Interaction of the PBD Dimer SJG-136 with Cognate Sequences of Oncogenic Transcription Factors

Mantaj, Julia

Awarding institution:
King's College London

Interaction of the PBD Dimer SJG-136 with Cognate Sequences of Oncogenic Transcription Factors

Julia Mantaj

A thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Institute of Pharmaceutical Sciences

King's College London

2016

Supervisors:

Dr Khondaker Miraz Rahman

Professor David E. Thurston

For my father Roman Mantaj

Thank you for passing on to me your passion for science.

Acknowledgements

My utmost thanks go to my primary supervisor Dr Khondaker Miraz Rahman and my secondary supervisor Professor David E. Thurston who both have been strongly involved and interested in my project. I am hugely thankful to both of my supervisors for always having an open mind towards my own ideas during my PhD and for our great working relationship. I am extremely grateful to Dr Khondaker Miraz Rahman for his excellent guidance, encouragement and support he gave me during the last four years. He has been a great supervisor. I am also deeply thankful to Professor David E. Thurston for his exemplary guidance, support and the valuable input he gave me during my project. In particular, I greatly appreciate all his hard work, support and guidance with my publications.

I would like to thank Dr Richard Parsons for all his help with the biology aspects of my project and for giving me the opportunity to learn new biological techniques. I was meant to spend only a few weeks in his lab but stayed for one year as I have enjoyed it so much. Thanks also to Dr Parson's former PhD student Dr Martin Thomas for accepting me from day one and treating me like a member of their group.

I am extremely thankful to Dr Paul Jackson, who has carried out the molecular dynamics simulation, for all his hard work and involvement in my project. But more importantly, I am thankful for our friendship. Thank you for having always your door open for me, for your motivating and encouraging words and for making me laugh.

I am thankful to all the members of the Thurston and Rahman research group for making my PhD enjoyable.

I would like to thank my friends (Magda, Helene and Kasia) for all their support and motivation over the last few years and particularly during my writing-up period. Thanks also to Dr Abhinav Kumar for all the encouraging chats and for providing me with delicious home-made dinners while I was writing.

Finally, I am hugely thankful to the most important people in my life, my brother and his family, my mother Erna and my father Roman for all the love, warmth and support they gave me throughout my life. I would not have achieved all of this without you, and I hope that I have made you proud.

Abbreviations

A	Adenine
Å	Angstrom
ADC	Antibody-Drug Conjugate
AKT	Protein Kinase B
AP-1	Activating Protein-1
ATP	Adenosine Triphosphate
AUC	Area Under the Curve
Bcl-2	B-cell like lymphoma-2
Bcl-xL	B-cell lymphoma-extra Large
bp	Base pairs
BPTI	Bovine Pancreatic Trypsin Inhibitor
C	Cytosine
CD	Circular Dichroism
CD20	B-Lymphocyte antigen CD20
CDK	Cyclin-Dependent-Kinase
cDNA	Complementary Deoxyribonucleic Acid
CLL	Chronic Lymphocytic Leukaemia
C_{max}	Maximum plasma concentration
CREB	Cyclic Adenosine Monophosphate Response Element Binding Protein
GM-CFU	Granulocyte-Macrophage Colony Forming Unit
CML	Chronic Myeloid Leukaemia
c-myc	Avian myelocytomatosis virus oncogene cellular homolog
CR	Complete Response
Da	Dalton
ΔCt	Increment in threshold cycle
ΔT_m	Increment in melting temperature
DHFR	Dihydrofolate Reductase
DLT	Dose Limited Toxicity
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribose Nucleoside Triphosphate
EGF	Epidermal Growth Factor
EGFR	Epidermal Growth Factor Receptor
EGR-1	Early Growth Response-1
Elk-1	E twenty six-like transcription factor-1
ERK	Extracellular-signal-Regulated Kinase
FAM	6-Carboxyfluorescin
FGF	Fibroblast Growth Factor
FRET	Fluorescence Resonance Energy Transfer
G	Guanine
GAPDH	Glyceraldehyde 3-Phosphate Dehydrogenase
G-CSF	Granulocyte-Colony Stimulating Factor
GI_{50}	Half maximal proliferation inhibitory concentration
GPU	Graphics Processing Unit
HER-2	Human Epidermal Growth factor-2
HIV	Human Immunodeficiency Virus
HPLC	High Performance Liquid Chromatography
HRP	Horseradish Peroxidase

HRAS	Harvey Rat Sarcoma Viral Oncogene Homolog
I	Inosine
IC_{50}	Half maximal inhibitory concentration
ICL	Interstrand Cross-Link
IgG1	Immunoglobulin G1
IKK	IκB Kinase complex
IL	Interleukin
JAK	Janus Kinase
JNK	C-Jun N-terminal Kinase
kDa	Kilo Dalton
KRAS	Kirsten Rat Sarcoma Viral Oncogene Homolog
LC	Liquid Chromatography
LC_{50}	Half maximal lethal concentration
LPS	Bacterial Lipopolysaccharide
MALDI-TOF	Matrix-Associated Laser Desorption/Ionisation Time-Of-Flight
MAPK	Mitogen-Activated Protein Kinase
MD	Molecular Dynamics
mdeg	Millidegrees
MED	Minimum-Effective Dose
min	Minute
MM	Molecular Modeling
mRNA	Messenger Ribonucleic Acid
MS	Mass Spectrometry
MTD	Maximum-Tolerated Dose
mTOR	Mammalian Target Of Rapamycin
MTT	3-(4,5-Dimethylthiazol-2yl)-2,5-diphenyltetrazol
m/z	Mass to charge ratio
NCI	National Cancer Institute
NEMO	NF-κB Essential Modulator
NF-κB	Nuclear Factor κappa B
NF-Y	Nuclear Transcription Factor Y
NHL	Non-Hodgkin's Lymphoma
NIK	NF-κB Inducing Kinase
NMR	Nuclear Magnetic Resonance
NNMT	Nicotinamide N-Methyltransferase
NSCLC	Non-Small Cell Lung Cancer
p-values	Probability values
p53	Tumour protein 53
PBD	Pyrrolobenzodiazepine
PCR	Polymerase Chain Reaction
PD	Pharmacodynamic
PDGF	Platelet-Derived Growth Factor
PI3K	Phosphatidylinositol-3-OH Kinase
PKA	Protein Kinase-A
pSTAT1	Phosphorylated STAT1
pSTAT3	Phosphorylated STAT3
PR	Partial Response
PTEN	Phosphatase and Tensin Homolog
Pu	Purine
Py	Pyrimidine
qPCR	Quantitative Polymerase Chain Reaction
(*R*)	Right-handed (rectus) stereochemistry

Rb	Retinoblastoma protein
RNA	Ribonucleic Acid
ROS	Reactive Oxygen Species
RP	Reverse Phase
RT	Retention Time
(*S*)	Left-handed (sinister) stereochemistry
SAR	Structure Activity Relationships
SD	Stable Disease
SD	Standard Deviation
SDS-PAGE	Sodium Dodecyl Sulfate- Polyacrylamide Gel Electrophoresis
SH-2	Src Homology-2
SOCS	Suppressors Of Cytokine Signalling
STAT	Signal Transducer and Activator of Transcription
T	Thymine
$t_{1/2}$	Half-life
TAMRA	6-Carboxytetramethylrhodamine
TF	Transcription Factor
TGI	Tumour Growth Inhibition
T_m	Melting temperature
TNF-α	Tumour Necrosis Factor alpha
tRNA	Transfer Ribonucleic Acid
tSTAT1	Total STAT1
tSTAT3	Total STAT3
U	Uracil
Ub	Ubiquitin
UV	Ultraviolet
VEGF	Vascular Epidermal Growth Factor
VEGFR	Vascular Epidermal Growth Factor Receptor
WB	Western Blotting

Abstract

The pyrrolo[2,1-*c*][1,4]benzodiazepines (PBDs) are sequence-selective DNA minor-groove interacting agents. They have a chiral centre at their C11a(*S*)-position which provides an appropriate three-dimensional shape for them to fit securely within the DNA minor-groove. They also possess a "soft" electrophilic imine moiety at their N10-C11 position which can form an aminal bond between their C11-position and the C2-NH_2 group of a guanine base only when the molecule is secure within the minor groove. The PBD dimer SJG-136 has currently reached Phase II clinical trials in ovarian cancer and leukaemia in the UK and the USA, respectively. More recently, PBD dimer analogues are being attached to tumour-targeting antibodies to create Antibody-Drug Conjugates (ADCs), some which are now in early clinical trials with many others in pre-clinical development.

Transcription factors (TFs) are sequence-specific DNA-binding proteins that bind to consensus DNA sequences, thereby controlling transcription. TFs regulate processes such as cell differentiation, proliferation and apoptosis. The interaction of a small-molecule with the consensus DNA sequences of TFs can prevent a TF from recognising its cognate sequence, thereby inhibiting the expression of genes critical for the survival and proliferation of cancer cells. The cognate sequences of a number of oncogenic TFs are unusually rich in GC sequences and provide ideal alkylation substrates (both cross-link and mono-alkylation modes) for both PBD dimers (*e.g.*, SJG-136) and monomers (*e.g.*, GWL-78 and KMR-28-39). There is growing evidence that PBDs exert their pharmacological effect through TF inhibition in addition to the arrest of the replication fork, DNA strand breakage, and inhibition of enzymes including endonucleases and RNA polymerases. For this reason, there is now interest in using PBDs as the basis for a small-molecule strategy to target specific DNA sequences for TF inhibition as a novel anticancer therapy.

This project initially involved the application of a reversed-phase HPLC/MS method as a tool to evaluate the interaction of DNA-binding PBD molecules with oligonucleotides of varying lengths and sequences. Using custom oligonucleotides containing the cognate sequences of NF-κB, EGR-1, AP-1 and STAT3 transcription factors, all of which contain

ideal GC-rich binding sequences for the PBD dimer SJG-136, the HPLC/MS method was used to provide information on both kinetic and thermodynamic adduct formation. Interestingly, a significant difference in the rate and extent of adduct formation between the different cognate sequences was observed which may explain the differences in activity of SJG-136 in the various tumour cell lines. Furthermore, SJG-136 was found to form different types of adducts with the TF sequences studied indicating a preference for guanine binding sites within the individual sequences. In particular, the fastest reaction was observed with the AP-1 sequence (*i.e.*, 100% adduct formation, and only one type of adduct formed) which led to guanine/inosine replacement studies to establish the precise DNA-binding site of SJG-136. During this study, the ability of PBD monomers (*e.g.*, GWL-78) and dimers (*e.g.*, SJG-136) to bond to a terminal guanine was demonstrated for the first time.

Further cellular experiments, the Polymerase Chain Reaction (PCR) and Western blotting, were carried out to support the observations made during the biophysical evaluation of SJG-136 with the NF-κB, EGR-1, AP-1 and STAT3 sequences. In these studies, SJG-136 was shown to significantly down-regulate a number of AP-1- and STAT3-dependent genes in the human colon carcinoma cell line HT-29 and the human breast cancer cell line MDA-MB-231.

Overall, the findings in this project have added significantly to the knowledge of the mechanism of action of SJG-136 and related PBD compounds, and could be important for the correct interpretation of the results of future pre-clinical and clinical studies of molecules of this type.

Table of contents

Chapter 1: Introduction

1.1 Cancer

1.1.1 Cancer epidemiology

Cancer is a disease of chaos characterised by a breakdown of the existing biological order within the body[1]. Cancer is an immense worldwide health burden as it was the leading cause of death worldwide in 2008[2]. Various human cancers affect the population in every region of the world and at a socioeconomic level. In 2012, 14.1 million new cancer cases were diagnosed and 8.2 million cancer deaths were reported worldwide (around 13% of all deaths)[2]. Despite the overall decline of cancer death rates (20% from their peak in 1991), the mortality rate due to cancer is estimated to rise to over 13.1 million by 2030[2-3].

Incidence and mortality rates of various cancers vary depending on the gender in addition to the country[2-4]. For example, lung cancer is the leading cause of cancer deaths in men worldwide (with about 1.4 million deaths in 2008) followed by liver and stomach cancers. Lung cancer is also the most frequently diagnosed cancer in most parts of Eastern Europe and Asia followed by prostate and colorectal cancers (**Figure 1.1**). Among women, breast cancer is the leading cause of death followed by lung cancer. Breast cancer is also the most commonly diagnosed cancer in most countries (*e.g.*, Australia, Europe, America and Asia) followed by cervical and colorectal cancers (**Figure 1.1**).

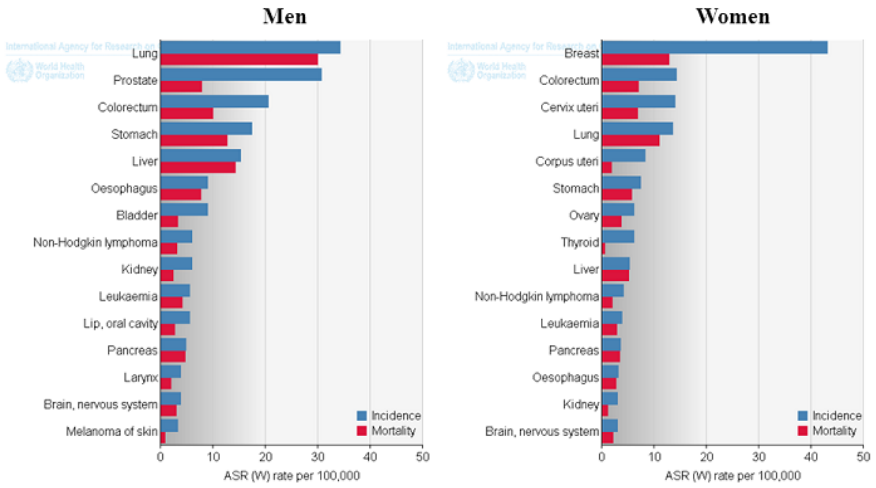


Figure 1.1: Estimated incidence and mortality rates in various cancers in men and women 2012. (Accessed on *http//globocan.iarc.fr*).

1.1.2 Causes of cancer

Cancer is caused by spontaneous mutations in genes. Therefore, it can also be referred to as a genetic disease. Alterations in the genomic sequences result in the ability of cancer cells to acquire highly abnormal phenotypes which disobey construction and maintenance rules of normal and intact tissues[1]. The vast majority of human cancers, estimated 90%, are associated with population exposure to environmental risk factors including tobacco-specific carcinogens (*e.g.*, polynuclear aromatic hydrocarbons)[5], dietary lifestyle[6], chemical agents (*e.g.*, asbestos[7], vinyl chloride[8] or aromatic amines[9]), bacteria[10] and viruses[11]. Moreover, hormones such as oestrogen and testosterone were reported to be implicated in carcinogenesis of hormone-associated cancers, including breast, endometrium, ovary and prostate by promoting cell proliferation[12]. Furthermore, aberrant oncogene activity provokes increased production of encoded proteins, such as cell surface receptors (*e.g.*, EGFR, VEGFR) that transmit extracellular signals into the inside of the cell as well as intracellular proteins participating down-stream in cell growth and division signalling pathways (*i.e.*, HRAS, KRAS or cyclin D1), thereby driving tumour cell division and proliferation[13]. Proteins encoded by oncogenes are usually transcription factors (*e.g.*, AP-1), chromatin remodelers, growth factors (*e.g.*, PDGF), growth factor receptors (*e.g.*, EGFR, VEGFR), signal transducers (*e.g.*, AKT) or apoptosis regulators (*e.g.*, Bcl-2, Bcl-xL)[14].

1.1.3 Cancer biology

The evolution of normal cells into neoplastic phenotypes is a complex multi-step process that normally proceeds over a period of decades. During this time, the tumour cells acquire a succession of various biological capabilities that are required for cancer cells to survive, proliferate and disseminate, and subsequently become a life-threatening disease. Hanahan and Weinberg proposed six key behaviours of cancer cells that rationalise the complexities of malignant diseases, also termed The 'Hallmarks of Cancer'[15] (**Figure 1.2**).

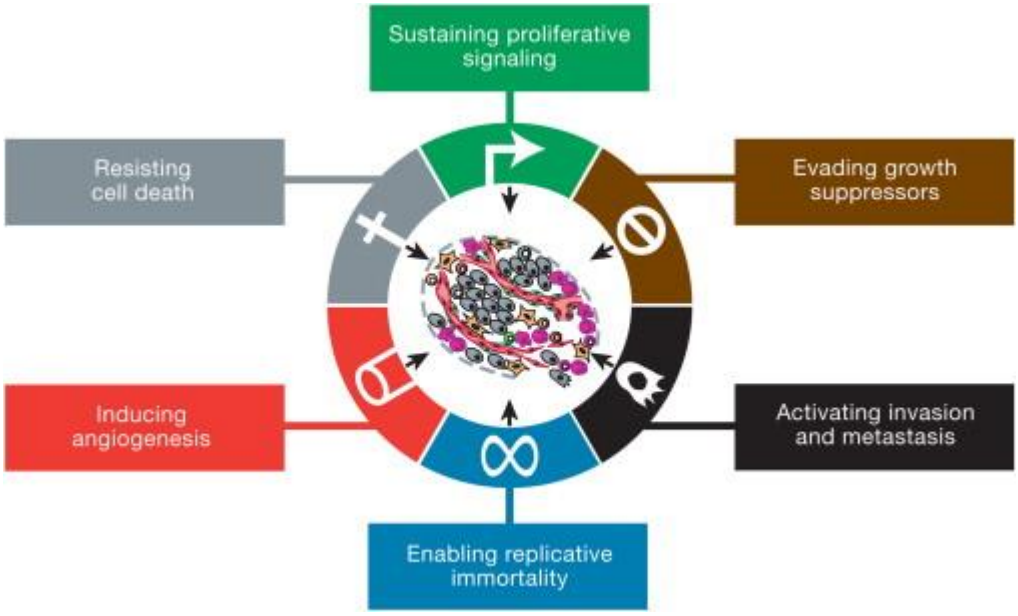


Figure 1.2: The six hallmarks of cancer. Taken from *D. Hanahan and R.A. Weinberg; Hallmarks of cancer: the next generation; Cell. 2011 Mar 4;144(5):646-74.*

In tumour cells, dysregulation of the mitogenic signalling by a number of various mechanisms enables the cancer cells to defy control, hence they progress through the cell cycle and increase in cell size[15]. In addition, mutations in genes encoding tumour suppressors (*e.g.*, *Rb*, *p53*) are frequently detected in various cancers and are considered to contribute to tumour development and progression.

Aberrant expression and regulation of pro- and anti-apoptotic proteins (*e.g.*, the Bcl-2 family[16]), which are fundamental for the control of apoptosis, promotes tumourigenesis. For instance, remarkably elevated levels of Bcl-2 are frequently observed in diffuse large B-cell lymphoma[17], acute myelogenous leukaemia[18], glioblastoma[19] and lung cancer[20]. The tumour suppressor p53 may also influence the regulation of apoptosis by interacting with pro- and anti-apoptotic proteins in response to irreparable DNA damage or chromosomal abnormalities[21]. Consequently, loss of p53 tumour suppressor function results in the disruption of the programmed cell death.

Moreover, tumour cells are characterised by chronically activated angiogenesis through disruption of the pro- and anti-angiogenic signals, commonly *via* the overexpression of VEGF[22]. The angiogenic switch enables them to continually form and spread new blood vessels that support the expansion of the neoplastic growth[23].

Metastasis is defined as the spread of a tumour cell from the primary neoplasm to new distinct anatomical sites through the blood and lymphatic system[1]. Despite significant progress in cancer diagnosis, surgical, local and systemic therapies, metastasis is the most threatening aspect of cancer due to its resistance to conventional treatment. 90% of all cancer deaths are caused by metastasis and this highlights its relevance.

1.1.4 Treatment of cancer

The conventional therapies for the treatment of human cancers are aggressive strategies as they have only one main goal, to kill rapidly dividing cells. The traditional eradication of malignancies consists of surgery, chemotherapy and radiotherapy. Applied since 1940, the 'old-school' therapeutic forms proved to be highly effective in prolonging the survival of cancer patients and eradicating a variety of tumours with curative outcomes. Despite significant improvements in understanding the insights of disease pathogenesis and remarkable progress in the discovery of novel targeted therapeutic options, the conventional methods remain, to date, the mainstay in the treatment of various malignancies in the clinic and represent the backbone in cancer treatment plans.

1.1.4.1 Surgery

Before the era of chemotherapy in 1940s, surgery was the only available option for the treatment of cancer. The principle of surgery is resection of the primary tumour, which also helps to reduce the risk of metastatic relapse. Of utmost importance is the fact that surgical treatment is not able to eradicate disseminated and metastatic tumour cells. With this consideration in mind, development of cytotoxic drugs became inevitable and emerged onwards with the ultimate goal to eradicate primary tumours as well as metastatic cells.

1.1.4.2 Chemotherapy

Chemotherapy applies highly effective and cytotoxic compounds for the treatment of human malignancies, including nitrogen mustard analogues, antimetabolites, alkylating agents, mitosis inhibitors and topoisomerase inhibitors. While having distinct mechanisms of action, the biological outcome is similar. Conventional chemotherapeutic drugs induce DNA damage that tumour cells cannot survive, therefore culminating in apoptosis. The serious side effects, such as peripheral neuropathy, kidney, liver and cardiac toxicity in addition to myelosuppression become apparent as these agents destroy rapidly dividing cells without distinguishing between normal healthy cells and malignant cells[24].

1.1.4.2.1 Antimetabolites

Antimetabolites (**Figure 1.3**) are chemical analogues of essential metabolites that interfere in various ways with the normal functioning of specific metabolites or enzymes produced by normal and cancerous cells. Small changes in crucial metabolites that are needed by cells for replication and proliferation inhibit cell growth.

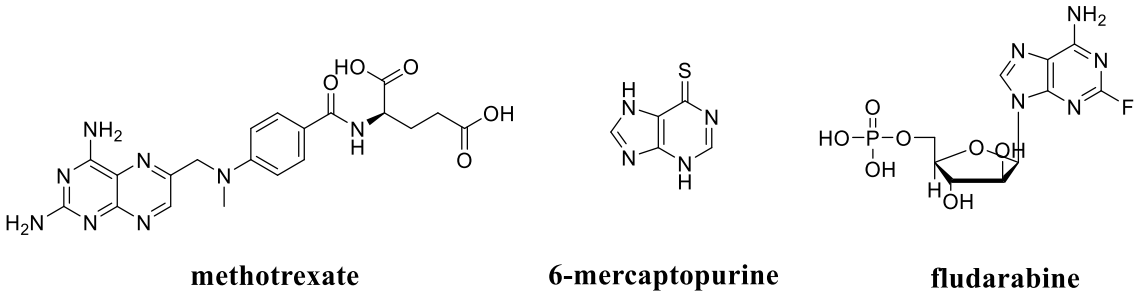


Figure 1.3: Chemical structures of the antimetabolites methotrexate, 6-mercaptopurine and fludarabine.

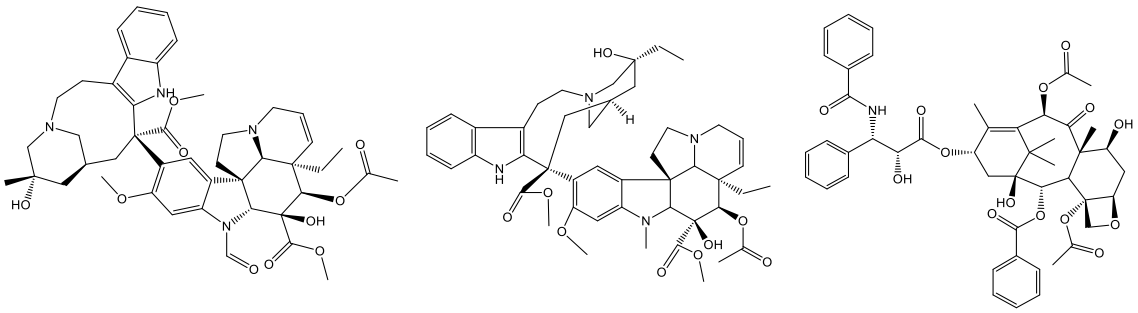
Antifolate drugs selectively inhibit the synthesis of thymidylate and purines by interfering with the dihydrofolate reductase (DHFR)[25]. The most striking example is methotrexate (**Figure 1.3**) which is used as single agent for the treatment of breast, ovarian, bladder and head and neck cancers[24].

Furthermore, purine analogues, such as 6-mercaptopurine and fludarabine (**Figure 1.3**) interfere with nucleoside biosynthesis, consequently creating irreparable DNA residues. Purine analogues are potential drugs in the treatment of leukaemia and Non-Hodgkin's lymphoma[1].

1.1.4.2.2 Mitosis inhibitors

Inhibition of the cell cycle can be achieved by interfering with the mitosis process. Disruption of the microtubule assembly (critical for chromosomal transport within the cell) results in an interruption to cell division and proliferation.

Vincristine and vinblastine are natural products derived from *Vinca rosea* (**Figure 1.4**)[26]. They intervene in the assembly of microtubules through interaction with the microtubule organising centres. Vinca-alkaloids are used in the clinic for the treatment of lymphomas, non-small cell lung cancer, breast and head and neck carcinomas[27].



vincristine **vinblastine** **paclitaxel**

Figure 1.4: Chemical structures of the mitosis inhibitors vincristine, vinblastine and paclitaxel.

In contrast, paclitaxel (**Figure 1.4**), a natural product of *Taxus brevifolia* interrupts mitosis by blocking the breakdown of microtubules at the end of the cell division process. It possesses a potent therapeutic activity on a variety of commonly occurring malignancies, including lung, ovary, breast and head and neck cancers[28].

1.1.4.3 Radiotherapy

Radiation therapy became an important clinical tool for the treatment of cancer after 1960. Direct exposure to electromagnetic radiation leads to burn damage of normal and malignant tissue thereby controlling local and regional tumours. Furthermore, radiotherapy is a valuable option for reducing post-surgical relapse. Alternatively, radioactive compounds can be administered intravenously as part of a pharmaceutical drug or coupled to a monoclonal antibody.

1.1.4.4 The targeted treatment of cancer

Significant progress in genomics, understanding the biology of malignancies and improved technology, has recently led anti-cancer drug development in a completely new direction. Targeted therapies are rationally designed to inhibit the activity of specific molecular targets within cancer cells, which are usually proteins whose aberrant function contributes to the formation and maintenance of tumours. Specific examples of these targets are growth factors, signalling proteins, cell-cycle proteins, regulators of apoptosis and molecules that promote angiogenesis[29].

For instance, anti-estrogenes such as tamoxifen, anastrozole and exemestane are used for the treatment of breast cancer patients[30].

The design of small-molecules aimed to selectively inhibit molecular targets that fundamentally contribute to carcinogenesis, tumour maintenance and metastasis. The most popular small-molecule inhibitors are imatinib mesylate (Gleevec®) (**Figure 1.5**) for the treatment of chronic myeloid leukaemia (CML), the epidermal growth factor receptor (EGFR) tyrosine kinase inhibitor gefitinib (Iressa®) (**Figure 1.5**), successfully used in clinics for the treatment of non-small cell lung carcinoma (NSCLC) and the vascular epidermal growth receptor (VEGFR) kinase inhibitor sorafenib (Nexavar®) (**Figure 1.5**)[31].

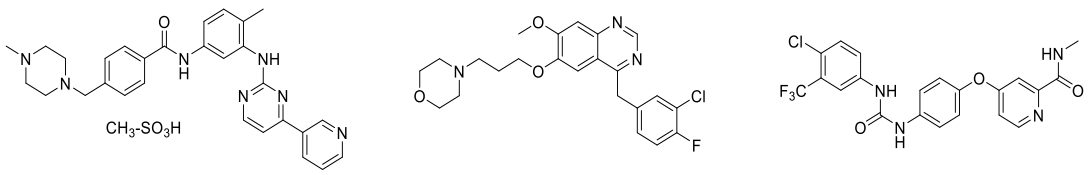


Figure 1.5: Chemical structures of the small-molecules imatinib, gefitinib and sorafenib.

Furthermore, monoclonal antibodies have become significant and indispensable tools for the treatment of human malignancies[32]. The introduction of rituximab for therapeutic use has revolutionised antibody design and development.

Rituximab (Mabtera®, Rituxan®), a chimeric monoclonal antibody, binds with high affinity to the transmembrane protein CD20 that is present on the surface of almost all (90%) malignant B cells of Non-Hodgkin's lymphoma (NHL)[33].

Trastuzumab (Herceptin®, Tykerb®), a humanised IgG1 monoclonal antibody, selectively interacts with the extracellular domain of the human epidermal growth factor receptor 2 (HER-2)[34]. HER-2 possesses a central position in tumorigenesis as it governs cell proliferation, survival and differentiation through interconnected signal transduction implying the activation of the PI3K-AKT and the MAPK signalling cascades[35].

1.2 Transcription factors

1.2.1 Overview

Throughout an organism's life, individual biological processes, such as cell development, differentiation and cell growth are governed by highly structured and coordinated regulatory networks on the transcriptional level[36]. Transcription factors (TFs) are sequence-specific DNA-binding proteins which are induced in response to particular external and internal stimuli arising from environmental needs of cells, controlling the expression of specific genes. TFs exhibit a pivotal role in regulating either positively or negatively gene transcription and deciding whether the primary gene transcript will result in the synthesis of the corresponding protein[36b]. Some TFs are constitutively and ubiquitously expressed over the lifespan of an organism. However, the expression of a number of TFs varies dramatically between characteristic tissues and corresponds with their regulatory requirements[36a]. Advances in human genome sequencing revealed the presence of 2,000-3,000 proteins that are presumed to contain specific DNA-binding domains.

1.2.2 Role of transcription factors in cancer

Given the pivotal role of TFs in a wide range of biological cellular processes, it has become increasingly clear that transcriptional misregulation is highly likely to be associated with human diseases[37]. Incorrect regulation and aberrant activity of regulatory TFs due to mutations are strongly linked with disorders, including inflammation, viral infections, autoimmune diseases and human malignancies. Uncontrolled TF activation continuously drives cell division, proliferation and survival through enhanced expression of growth stimulating gene products, thereby crucially contributing to tumour development and progression. Accordingly, TFs represent a valuable target for the design and the development of anti-tumour strategies which include inhibition of their functional domains or their post-translational modification with the goal to terminate their uncontrolled transcriptional activity.

1.2.3 Transcription factors used in this study

1.2.3.1 Nuclear Factor-κB (NF-κB)

The NF-κB proteins belong to a family of structurally related, rapidly acting TFs that play a fundamental role in regulating a diverse range of physiological processes, including the innate-adaptive immune responses, inflammation, cell differentiation and proliferation as well as apoptosis[38].

The NF-κB TFs exist as dimers composed of two of five possible subunits. These subunits may be divided into two subclasses; the Rel proteins RelA (p65), RelB and c-Rel (which contain the transcription activation domain, TAD), and the p50/105 and p52/100 members (which lack the TAD required for transcriptional activity)[39]. In order to exhibit transcriptional ability members of this second class require association with Rel proteins. The NF-κB proteins bind as homodimers or heterodimers to 9-10 base pair DNA sites (κB sites) in the promoters or enhancer regions of target genes recognising GGGRNYYYCC sequences (R = purine; Y = pyrimidine; N = any base), hence modulating the expression of specific genes (**Figure 1.6**)[40].

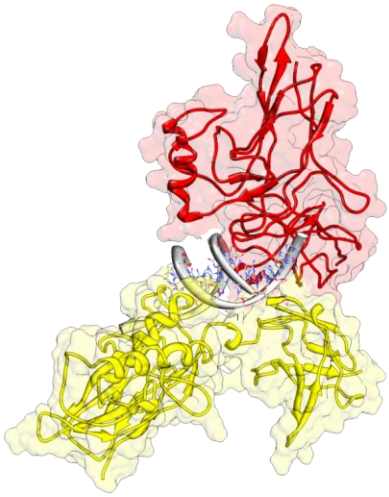


Figure 1.6: Molecular model of the NF-κB p50-p65 heterodimer (p50 in yellow and p65 in red) interacting with its DNA consensus sequence 5'-GGGGACTTTCC-3' (grey) (PDB ID: 1VKX)[41].

The NF-κB TFs are not synthesised *de novo*. They are present in the cytosol in a latent and inactive state bound to inhibitory IκB proteins (*e.g.*, IκBα, IκBβ, IκBγ or IκBε)[42]. They are activated *via* two distinct signalling pathways, namely the canonical (classical) or the non-canonical (alternative) pathway (**Figure 1.7**)[43].

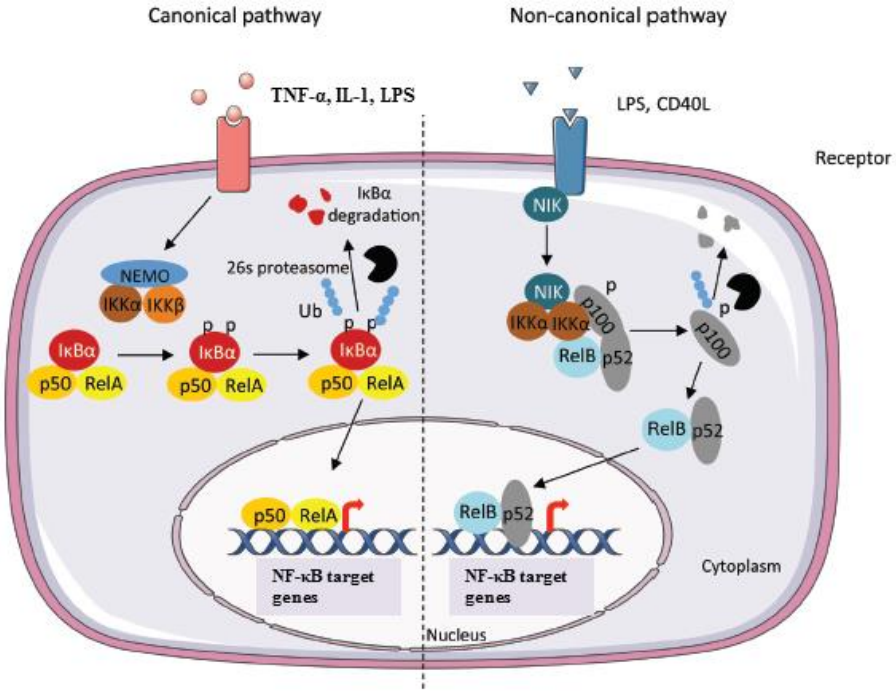


Figure 1.7: Canonical and non-canonical pathway of NF-κB activation. Adapted from *E. Viennois, F. Chen, D. Merlin; NF-κB pathway in colitis-associated cancers; Transl Gastrointest Cancer 2012;2(1); DOI:10.3978/j.issn.2224-4778.2012.11.01*.

Extracellular stimuli, such as viruses, bacterial lipopolysaccharide (LPS), oxidative stress, chemical and pharmacological agents, pro-inflammatory cytokines, tumour necrosis factor α (TNF-α) and UV radiation induce activation of the IκB kinase complex (IKK) comprising of IKKα, IKKβ and IKKγ (NEMO) members[44] which in turn phosphorylate the inhibitory IκB proteins leading to their ubiquitination and proteasomal degradation. As a consequence, the unmasked NF-κB proteins are able to translocate to the nucleus where they influence the expression of specific target genes either alone or in combination with other TF families (*i.e.*, AP-1 or STAT).

Furthermore, *via* the canonical pathway, NF-κB encodes genes for IκB *de novo* synthesis. That creates a negative feedback loop. Newly synthesised IκB proteins migrate to the nucleus, remove associated NF-κB from the DNA and export it back to the cytoplasm in order to restore the original inactive state.

The canonical pathway is responsible for the control of the innate immunity and inflammation[43], whereas the non-canonical pathway occurs mainly during the development, organisation and function of secondary lymphoid organs and in B cell differentiation and survival[43].

Recent studies have provided strong evidence for aberrant activation of the NF-κB signalling cascade in both solid tumours (*e.g.*, breast, cervical, prostate, renal, lung, colon, liver and pancreatic tumours) and haematological malignancies (*i.e.*, leukaemias and lymphomas)[45]. Genetic alterations, including gene mutations or amplifications in key components of the NF-κB signalling pathway (*e.g.*, RelA, RelB, c-Rel, IKKs and IκBs), over-expression of target genes and their products (*i.e.*, anti-apoptotic members of the Bcl-2 family, survivin, cyclin D1) and chromosomal translocations are presumed to contribute to cancer cell proliferation, growth, angiogenesis, metastasis and resistance to apoptosis[46].

When viewed from this perspective, the inhibition of major components of the NF-κB pathway is a considerable and powerful strategy for the treatment of neoplastic diseases.

1.2.3.2 Early Growth Response-1 (EGR-1)

EGR-1 (**Figure 1.8**) belongs to the EGR family of TFs which are implicated in a wide range of normal cellular processes, such as cell growth, proliferation, differentiation, survival and apoptosis[47]. The EGR family is composed of four members, EGR-1, EGR-2, EGR-3 and EGR-4 of which all are characterised by a highly conserved DNA-binding domain containing three zinc-finger motifs (**Figure 1.8**)[48]. These fingers recognise consensus DNA target sequences of 5'-GCG GGG GCG-3' with each finger spanning three bases (**Figure 1.8**)[49]. EGR-1 is rapidly and transiently expressed in response to a variety of stimuli, including growth factors, cytokines, environmental and mechanical stress[50]. Bound to promoter and enhancer regions in proximity to specific target genes,

they influence the expression of hundreds of genes by a mechanism that is still poorly defined.

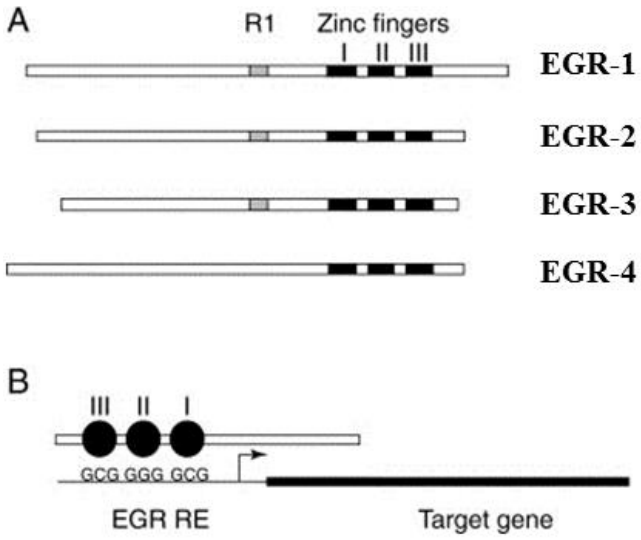


Figure 1.8: The EGR TF family and its cognate response element. (**A**) The four EGR-family members containing the three zinc-finger DNA-binding domains (black). R1 represents the repression domain. (**B**) An EGR-family member is binding to the consensus EGR sequence located in close proximity to the target gene. Each zinc finger (black circle) binds to a three-nucleotide segment. Adapted from *K. J. O'Donovan, W. G. Tourtellotte, J. Millbrandt, J. M. Baraban; Trends in Neurosiences; Vol 22, Issue 4, 1 April 1999, Pages 167–173; doi:10.1016/S0166-2236(98)01343-5.*

In the context of human malignancies, EGR-1 is considered as a tumour suppressor by promoting the expression of the major tumour suppressor genes *p53, Rb* and *PTEN* and the pro-apoptotic gene *c-Jun*[51]. Decreased levels of EGR-1 expression were observed in various human tumour types (*e.g.*, glioblastomas, lymphomas, non-small cell lung tumours, breast and colon cancers), therefore strongly linked with disease development and progression[52].

However, sufficient evidence has emerged of a role for EGR-1 in proto-oncogene function, contributing to malignancy. EGR-1 over-activity has been shown to underpin the pathophysiology of various human tumours, including prostate, gastric cancers and diffuse large B-cell lymphoma[51]. Hence, inhibition of EGR-1 may be a possible anti-tumour target for human neoplastic diseases owing aberrant EGR-1 expression.

1.2.3.3 Activating Protein-1 (AP-1)

The activating protein-1 (AP-1) proteins are a family of dimeric regulatory TFs, comprising four subfamilies; the Jun (c-Jun, JunB and JunD), the Fos (c-Fos, FosB, Fra1 and Fra2), the activating TFs (ATF) proteins (ATF2, ATF3/LRF1 and B-ATF) and lastly, the Maf subfamily members (c-Maf, MafA, MafB, MafG/F/K and Nrl)[53]. The proteins of the AP-1 family possess two common peculiarities; the leucine-zipper (bZIP) domains which are essential for dimerisation and the basic sequence which is required for the DNA-binding to AP-1 consensus motifs in the promoter and enhancer segments of numerous mammalian genes[54].

Jun and ATF proteins can homodimerise and heterodimerise, whereas members of the Fos subfamily can only exhibit transcriptional activity by forming heterodimers with the Jun proteins[55]. Jun-Jun and Jun-Fos dimers associate with the highest affinity to the response element 5'-TGA (C/G) TCA-3' (**Figure 1.9**)[56]. Jun-ATF heterodimers demonstrated a preference for the cognate 5'-TGA CG TCA-3' sequence[57]. However, due to the immense number of AP-1 proteins and consequently the broad combinatorial possibilities, they may bind to many other 'AP-1-like' sequences, depending on the AP-1 dimer composition. This huge combinatorial diversity also envisions the wide spectrum of genes which are regulated by AP-1 TFs[58].

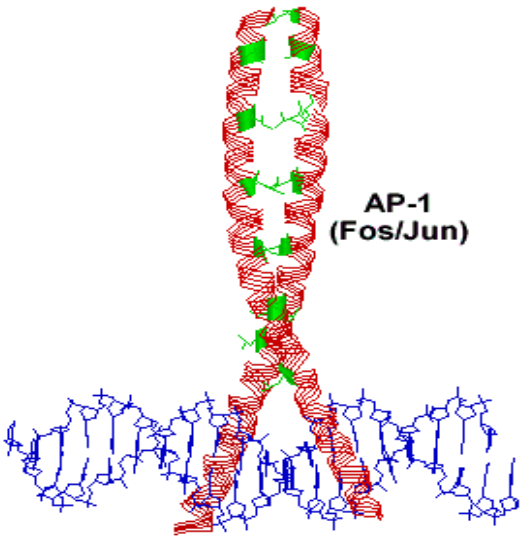


Figure 1.9: The structure of the heterodimer Fos-Jun AP-1 complex bound to its consensus target sequence on DNA. Taken from *http://www.web-books.com/MoBio/Free/Ch4F5.htm*.

The AP-1 proteins are implicated in the regulation of diverse biological processes, including inflammation, cell proliferation, differentiation and survival, apoptosis, cellular migration and wound healing (**Figure 1.10**)[58-59]. The activation of AP-1 members is initiated by growth factors, hormones, cytokines, stress, reactive oxygen species as well UV radiation and subsequently transmitted down-stream *via* the mitogen-activated protein kinase (MAPK) signalling pathway[59b].

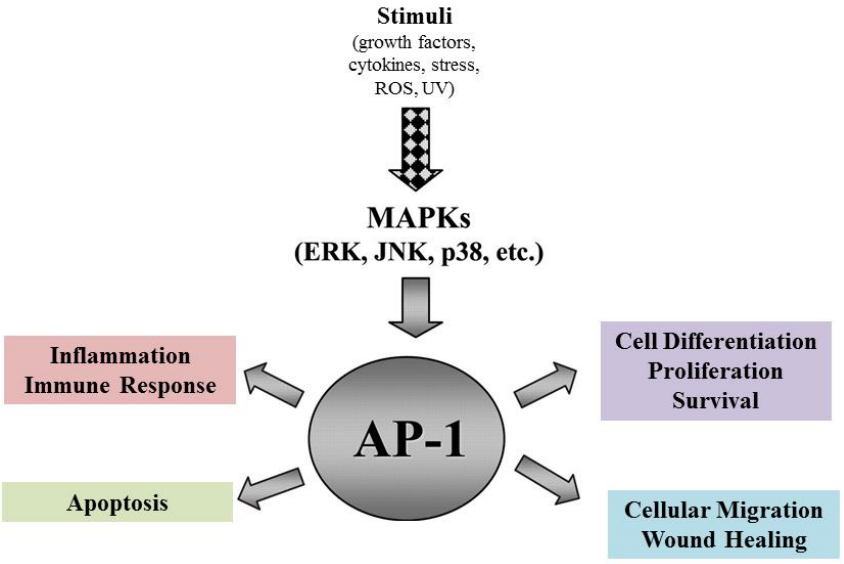


Figure 1.10: Activation of the TF AP-1 by various stimuli. Stimulation of different mitogen-activated protein kinase (MAPK) signalling pathways activates AP-1 resulting in the expression of a wide variety of genes participating in essential physiological processes. Adapted from *S.P.M. Reddy, B.T. Mossman; American Journal of Physiology - Lung Cellular and Molecular Physiology Published 1 December 2002 Vol. 283 no. 6, L1161-L1178 DOI:10.1152/ajplung.00140.2002.*

Elevated protein expression levels of c-Jun, JunB and Fos members, observed in a subgroup of human neoplasms, such as colorectal carcinomas and Hodgkin's lymphomas are strongly linked with tumour development and promotion[54, 60]. The importance of AP-1 TFs in human malignancies leads to the deduction that AP-1 may be a promising molecular target for anti-cancer therapies.

1.2.3.4 Signal Transducer and Activator of Transcription 3 (STAT3)

The signal transducer and activator of transcription 3 (STAT3) (**Figure 1.11**) belongs to the STAT family of cytoplasmatic proteins. Seven mammalian STAT proteins have been identified (STAT1, STAT2, STAT3, STAT4, STAT5a, STAT5b and STAT6) which all relay extracellular signals to the nucleus, thereby regulating the transcription of specific genes and orchestrating fundamental physiological processes[61].

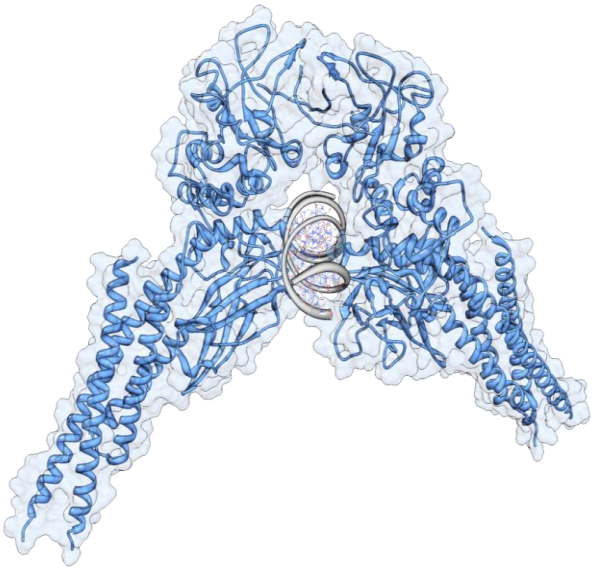


Figure 1.11: Molecular model of STAT3 homodimer (blue) interacting with the DNA sequence 5'-TGCATTTCCCGTAAATCT-3' (grey) (PDB ID: 1BG1)[62].

The STAT3 protein contains several substantial domains that are critical for its transcriptional activity. The DNA-binding domain is essential for the recognition of STAT3 consensus binding sequences on DNA. The SH-2 (Src homology-2) domain is required for the recognition of cell-surface receptor binding sites and lastly the C-terminal transactivation domain, necessary for transcription induction[63].

STAT3 is present in a latent and inactive state in the cytoplasm until it receives extracellular signals for activation. Growth factors (*e.g.*, EGF, PDGF, FGF or G-CSF) and cytokines (*i.e.*, IL-6, IL-10, IL-11 or oncostatin M) bind to their cognate cell-surface receptors and induce autophosphorylation and the activation of the Janus kinase (JAK) family (**Figure 1.12**). Following this, monomeric STAT3 molecules dock to the

phosphorylated tyrosine kinase complex through their SH-2 domain resulting in phosphorylation of STAT3 proteins in their C-terminal domain. This activation step allows STAT3 monomers to dimerise and translocate to the nucleus, where they bind to specific DNA response elements in the promoter regions of responsive target genes (**Figure 1.12**)[61, 63]. Alternatively, the mTOR (mammalian target of rapamycin) and MAPK signalling pathway lead also to phosphorylation and activation of STAT3[64].

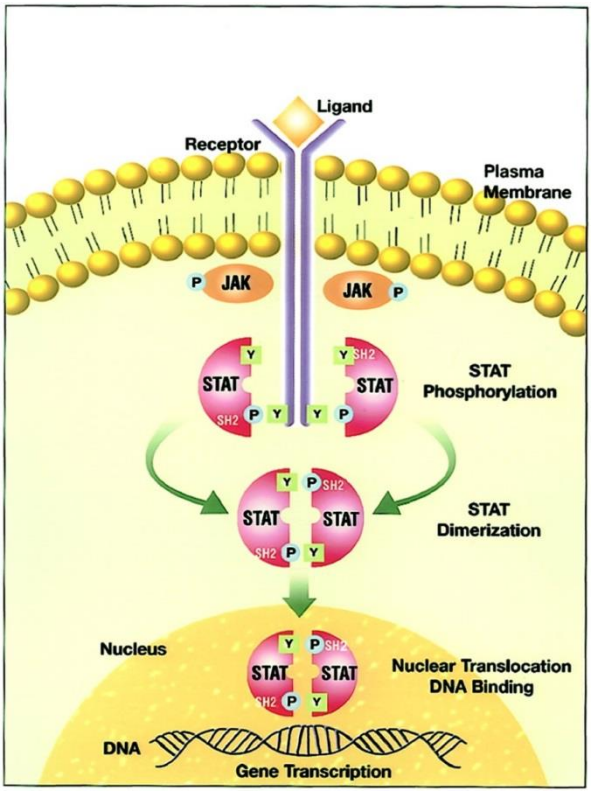


Figure 1.12: The JAK-STAT signal transduction pathway. Ligands activate JAKs that subsequently phosphorylates inactive STAT monomers. The STAT molecules dimerise and migrate to the nucleus, where they bind to specific response elements and induce gene transcription. Taken from *M. Benekli, M. R. Baer, H. Baumann and M. Wetzler; Signal transducer and activator of transcription proteins in leukaemias; Volume: 101 (8) p 2940 – 2954; DOI: http://dx.doi.org/10.1182/blood-2002-04-1204*.

Under normal conditions, the activity of STAT3 is endogenously controlled by specific regulatory proteins[61a, 61b, 63b]. For example, proteins of the SOCS (suppressors of cytokine signalling) family prevent STAT3 activation through directly binding to the

phosphorylated JAK. However, an increasing body of evidence suggests that elevated levels of STAT3 promote tumour cell proliferation, differentiation, angiogenesis and metastasis and protect neoplastic cells from apoptosis by regulating responsible genes, such as *Bcl-2*, *Bcl-xL*, *Fos*, *cyclin D*, *c-myc*, *VEGF*, *survivin* and *fascin*[61b, 63a, 65]. Increased STAT3 expression or activity is frequently observed in a wide range of human neoplasms, including leukaemias, lymphomas, multiple myelomas, breast, prostate, lung and ovarian cancers[61b, 65a, 66]. Interestingly, to date no naturally occurring mutations that lead to constitutive STAT3 activation have been detected indicating aberrant regulation or activation of key up-stream elements[61, 63b]. With this knowledge in hand, inhibition of the JAK/STAT3 signalling cascade may offer a novel strategy for the treatment of human malignancies.

1.3 DNA as a target for anti-tumour therapy

DNA is a valuable and attractive target for the development of anti-tumour therapies. Due to its inevitable presence in all living organisms, damage or disruption in DNA function results in destruction and/or death of cellular organisms.

DNA is composed of two anti-parallel arranged sugar-phosphate strands to which four different bases adenine (**A**), guanine (**G**), cytosine (**C**) and thymine (**T**) are linked, forming a double-stranded helix (**Figure 1.14a**). The two opposite running strands are stabilised through hydrogen bonds and base-stacking between the aromatic nucleobases[67]. Adenine pairs with thymine by forming two hydrogen bonds and guanine associates with cytosine by forming three hydrogen bonds (**Figure 1.14b**). The right-handed B-form is the most frequently observed conformation of DNA in which the base pairs are perpendicularly stacked to the double-helical axis.

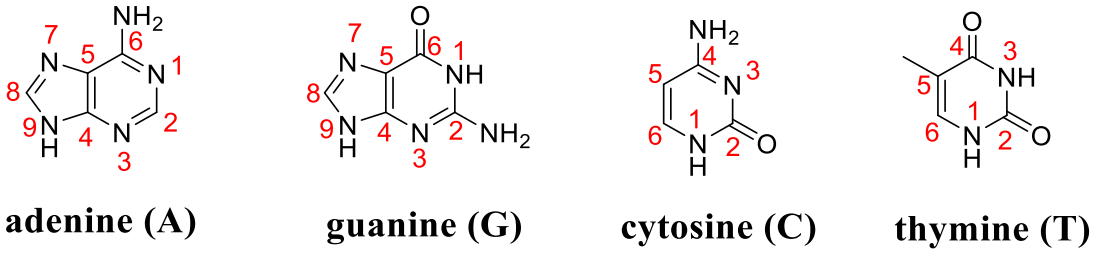


Figure 1.14: Nucleotide bases of DNA.

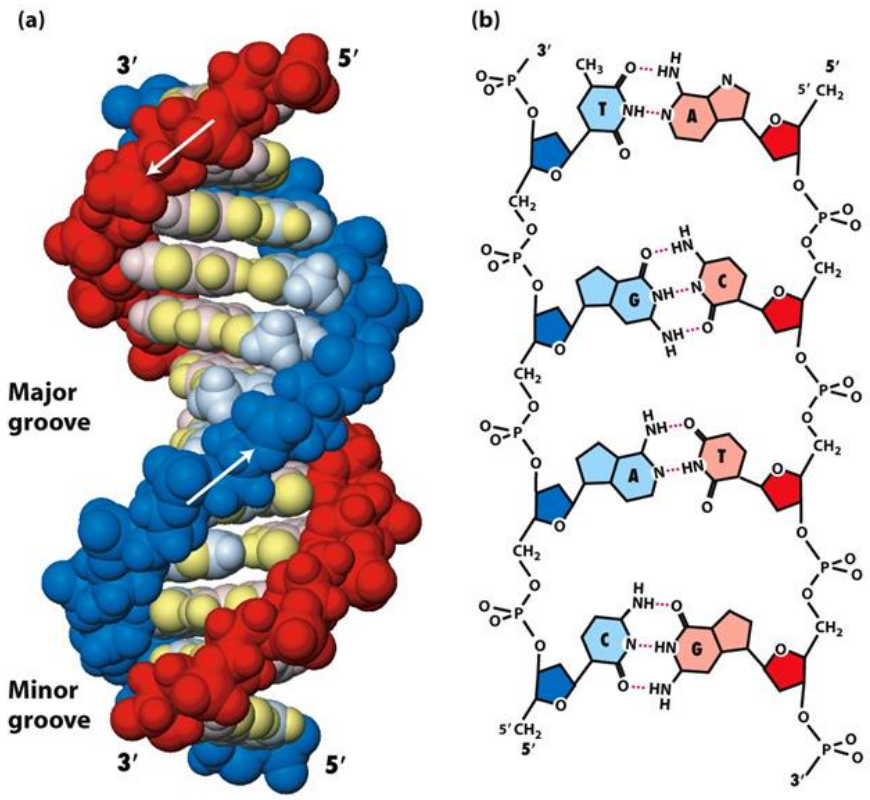


Figure 1.14: The structure of DNA. (**a**) B-form of DNA; (**b**) Base pairing between two DNA strands. T = thymine, A = adenine, C = cytosine, G = guanine. Taken from *Molecular Cell Biology, Sixth Edition © W.H. Freeman and Company.*

The adoption of the double helical conformation results in the formation of two grooves within the DNA helix, the major and the minor groove (**Figure 1.14a**). The major groove is wide and deep (22 Å) compared to the minor groove which is narrower and shallower (12 Å)[68]. This unique structure of the DNA double helix allows the appearance of additional hydrogen-bond donating and accepting groups located within the major and minor groove[69]. The major groove offers hydrogen bond donor groups at the C6-amino of adenine or the C4-amino of cytosine and hydrogen bond acceptor groups at the adenine-N7, thymine-O4, guanine-N7 and O6. The minor groove has hydrogen acceptor groups at adenine-N3, thymine-O2, guanine-N3 and cytosine-O2 and a hydrogen donor group at the 2-amino group of guanine.

For example, anti-tumour therapies can target DNA at the specific bases in order to interrupt crucial replication mechanisms or intercalate between the DNA strands and make them inoperable[70]. Moreover, inhibition of DNA-associated processes has gained

increased importance in anti-cancer drug development. For example, the transcription of specific genes which play a crucial role in the development and progression of human malignancies can be inhibited by small-molecule ligands[71] which block cognate sequences in double-stranded DNA, thereby preventing TFs to bind to their DNA target sequences[29, 31].

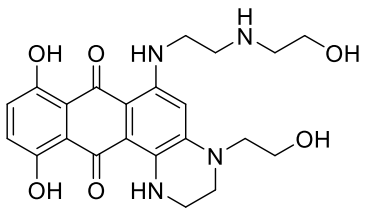
1.3.1 DNA-interactive agents

The vast majority of DNA-interactive agents directly target the DNA, such as alkylating agents and topoisomerase inhibitors, or interfere with the DNA synthesis (antimetabolites) (see also section 1.1.4.2.1)[69].

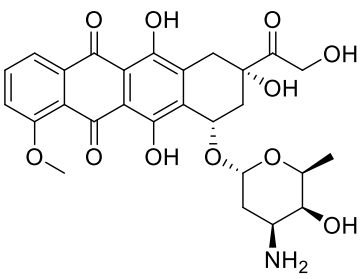
1.3.1.1 Topoisomerase inhibitors

Topoisomerases (topoisomerase I and topoisomerase II) are reversible nucleases that cleave super coiled DNA strands in order to ensure availability of DNA bases for DNA synthesis. Inhibition of topoisomerases through cytostatic drugs leads to the formation of irreparable DNA strand breaks forcing the cell to induce apoptosis. The activity of topoisomerases is elevated in many human tumours, therefore making it an attractive target for topoisomerase inhibitors.

Mitoxantrone (**Figure 1.15**) represents a powerful topoisomerase inhibitor in the treatment of acute myelogenous leukaemia, breast carcinoma and Non-Hodgkin's lymphoma[72].



mitoxantrone



doxorubicin

Figure 1.15: Chemical structures of the topoisomerase inhibitors mitoxantrone and doxorubicin.

The naturally occurring anthracycline doxorubicin (**Figure 1.15**) inhibits topoisomerase and is used in a wide range of human malignancies. In addition to the topoisomerase inhibiting potential, doxorubicin also has the ability to intercalate in the DNA, thereby blocking the DNA replication and to transform to free radicals resulting in double-strand breaks of the DNA[73].

1.3.1.2 Intercalators

Intercalating agents are polyaromatic planar molecules. Due to their flat shape they insert themselves between the DNA base pairs, thereby inhibiting essential processes, including transcription, DNA replication and repair[73-74].

The anthracyclines doxorubicin and daunorubicin (**Figure 1.16**) represent the best studied and widely used family of intercalating agents in the treatment of haematological malignancies, such as leukaemia, Hodgkin's diseases and lymphomas as well as solid tumours, including breast, ovarian and lung cancers[73]. These natural products, isolated from *Streptomyces* species exert their anti-tumour activity through a wide range of mechanisms, including DNA intercalation, inhibition of topoisomerase and formation of free radicals[73]. They are composed of a planar anthraquinone ring system linked to an amino-sugar through a glycoside bond. However, their toxicity (*e.g.*, cardiotoxicity) dramatically limits their use in clinic[75].

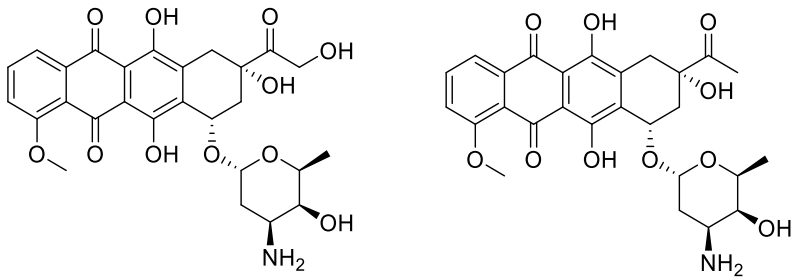


Figure 1.16: Chemical structures of the anthracyclines doxorubicin and daunorubicin.

1.3.1.3 Alkylating agents

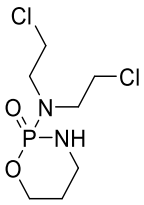
Alkylating agents exhibit their anti-tumour activity through the attachment of alkyl groups to guanine bases in DNA leading to DNA damage which cells cannot repair and, therefore, induce apoptosis[74, 76]. These DNA modifications trigger the DNA repair machinery[77]. Moreover, alkylating agents can bind to two guanines located either on the same or opposite DNA strand, thereby forming cross-linked adducts. This mechanism prevents separation of the DNA strands during replication and transcription[24, 77].

1.3.1.3.1 Cross-linking agents

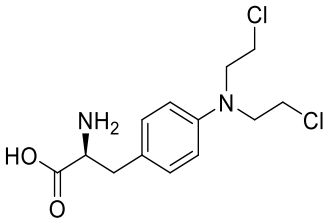
In addition to mono-adduct formation, cross-linking agents also react through their electrophilic moieties with DNA bases that are present on the same or opposite DNA strand, thereby generating the formation of cross-linked adducts within the DNA. Tumour cells lack the ability to repair the DNA damage and consequently induce apoptosis[78].

A) Nitrogen mustard analogues

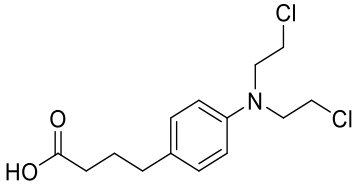
The most important examples of this group are cyclophosphamide, melphalan and chlorambucil (**Figure 1.17**). These anti-tumour agents are established standard components in the treatment of lymphomas, leukaemias and solid tumours[24, 79].



cyclophosphamide



melphalan

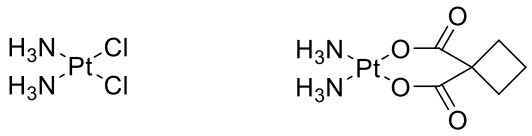


chlorambucil

Figure 1.17: Chemical structures of the nitrogen mustard analogues cyclophosphamide, melphalan and chlorambucil.

B) Cisplatin and carboplatin

Cisplatin and carboplatin are also members of the group of cross-linking agents (**Figure 1.18**). By forming intrastrand cross-links within the DNA they induce covalent DNA modifications that are not easy to repair for the tumour-repair machinery. They are used for the treatment of breast, ovarian and testicular cancers with 90-95% successful tumour eradication[1]. Cisplatin induces an intrastrand cross-link between the N7-N7 positions of guanines in the major groove[80].



cisplatin **carboplatin**

Figure 1.18: Chemical structures of cisplatin and carboplatin.

1.3.1.3.2 Minor groove binders

The traditional major groove binders such as the nitrogen mustards have long been used in the treatment for human malignancies. However, they demonstrate poor sequence specificity due to random adduct formation throughout the entire human genome. This unspecific binding is responsible for their cytotoxicity as they also kill rapidly dividing healthy cells, particularly of the bone marrow and gut epithelium[81]. Furthermore, studies have demonstrated that the N7 guanine adducts can easily be repaired by intracellular enzymes, resulting in resistance to chemotherapy[76]. Therefore, extensive effort has been made in order to design anti-cancer drugs that specifically bind within the DNA minor groove[74].

Minor groove binding agents are both naturally occurring and synthetic small-molecules that perfectly fit within the DNA minor groove because of their size and curvature[82]. They are composed of aromatic heterocycles and various linkers[83]. Studies have shown that the crescent shape of ligands significantly influences sequence specific ligand accommodation within the minor groove[84]. The interaction with DNA is mainly based on

non-covalent association caused through electrostatic forces, such as hydrogen bonding and van der Waals interactions. The majority of the minor groove binding agents exists in a positively charged form *in vivo*[85], therefore highly attracted to the electronegative potential of A/T sequences. The formed ligand/DNA adducts do not significantly alter the structure of DNA which makes them difficult to be recognised by intracellular DNA repair enzymes and eliminated. Moreover, agents that bind to the minor groove of DNA generally influence gene transcription by interfering with consensus sequences of many transcription factors. In biological studies, minor groove binding agents have demonstrated antibacterial, antiviral, antiprotozoal and anti-tumour activity[85].

Distamycin, Berenil and Hoechst 33258

Distamycin (**Figure 1.19**) is a natural product extracted from *Streptomyces* species[86]. It has a polyamide structure consisting of five-membered planar rings. The ligand shows a preference for AT regions by binding to four or five successive A-T base pairs within the minor groove. Hydrogen bonds are formed between the ligand and the O2 atoms of thymine and N3 atoms of adenines which help the molecule to accommodate within the minor groove.

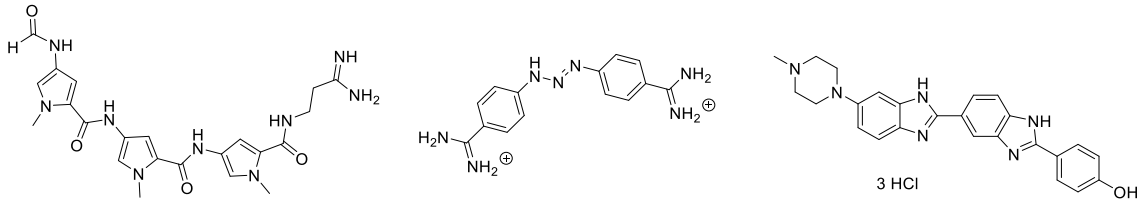


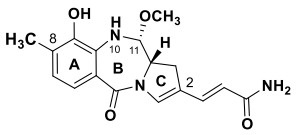
Figure 1.19: Chemical structures of the minor groove binding agents distamycin, berenil and Hoechst 33258.

Berenil (**Figure 1.19**) is a bis-phenylamidinium minor groove binding agent composed of two linked phenyl rings. The ligand binds preferentially to 5'-AAT-3' motifs where the amidinium groups form hydrogen bonds to the O2 of thymine or N3 adenine atoms.

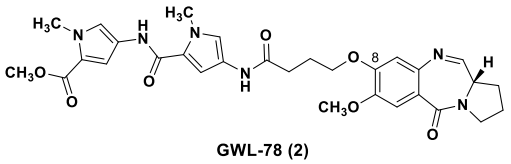
Another AT selective minor groove binder is Hoechst 33258 (**Figure 1.19**). Footprinting studies have shown that the benzimidazole-based ligand preferably binds to 5'-AATT-3' sequences[87]. However, due to its toxicity the agent has not completed clinical trials.

1.4 Pyrrolobenzodiazepines

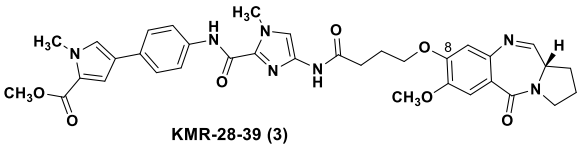
The pyrrolo[2,1-*c*][1,4]benzodiazepines (PBDs) are sequence-selective DNA minor groove binding agents based on the naturally occurring anthramycin family of anti-tumour antibiotics[83, 88], the best known member of which is anthramycin itself (**1**, **Figure 1.20**). The skeletal structure of the PBDs contains a substituted aromatic A-ring, a diazepine B-ring and a pyrrolidine C-ring[89], with an *S*-chiral centre at the C11a-position between the B and C rings. This provides a 3-dimensional shape perfectly crafted for the molecules to fit within the DNA minor groove[90] (**Figure 1.21**). They also possess an electrophilic imine moiety at the N10-C11 position within their B-ring which can form a covalent aminal linkage between their C11-carbon and the C2-NH_2 group of a guanine base (**Figure 1.21**), but only after the molecule is secure within the minor groove[91].



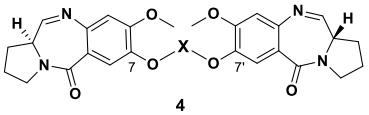
Anthramycin (1)



GWL-78 (2)

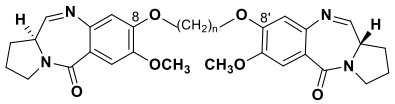


KMR-28-39 (3)



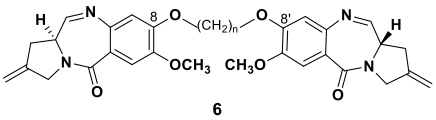
4

4a: X = $S(CH_2)_6S$
b: X = $O(CH_2)_2N(CH_3)(CH_2)_2O$
c: X = $HN(CH_2)_3N(CH_3)(CH_2)_3NH$



5

5a: n = 3 (**DSB-120**)
b: n = 4
c: n = 5
d: n = 6



6

6a: n = 3 (**SJG-136**)
b: n = 5 (**DRG-16**)

Figure 1.20: Structures of the naturally occurring anthramycin (**1**), the PBD C8-conjugates GWL-78 (**2**) and KMR-28-39 (**3**), and examples of C7/C7'-linked (**4**) and C8/C8'-linked (*e.g.*, DSB-120, **5a**, n = 3 and SJG-136, **6a**, n = 3) PBD dimers.

Since the discovery of anthramycin in 1965[92] a range of synthetic PBDs has been developed. For example, non-covalent minor groove binding components have been appended to the C8-position of the PBD A-ring (*e.g.*, GWL-78[89b], **2**, and KMR-28-39[90], **3**, **Figure 1.20**), and monomeric PBD units have been joined through their C7/C7' [93] (**4**) and C8/C8'-positions (*e.g.*, **5a-d** and **6a**,**b**) to afford PBD dimers (*e.g.*, SJG-136[94], **6a**, **Figure 1.20**). PBD units have also been joined through their C2/C2' and C8/C2' positions but the resulting dimers do not have the appropriate shape to fit into the DNA minor groove, and so have poor DNA-binding affinity and cytotoxicity[95].

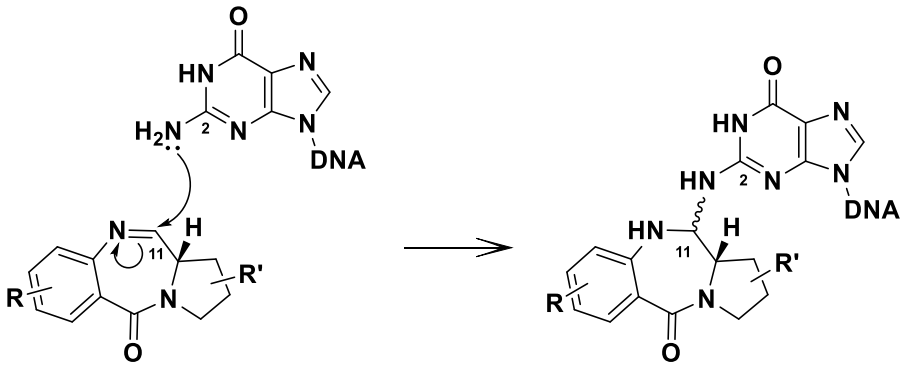


Figure 1.21: The mechanism of covalent binding of a PBD to DNA once it has located in a low energy position within the minor groove.

The unique structure of PBD dimers allows them to form interstrand or intrastrand DNA cross-links in addition to mono-alkylated adducts[96], thus resulting in greater DNA stabilisation[94, 97] compared to the monomeric PBDs. Due to the additional types of DNA adducts possible and the greater adduct stability, the PBD dimers generally have significantly greater cytotoxicity, anti-tumour activity and antibacterial activity compared to the natural, unsubstituted PBD monomers. One such agent, SJG-136, has reached Phase II clinical trials in ovarian cancer and leukaemia[98]. In addition, related PBD dimers are now being used as payloads for Antibody-Drug Conjugates (ADCs) due to their significant cytotoxicity[99].

The covalent binding of both monomeric and dimeric PBDs to DNA is thought to be a two-step process, the first involving recognition of a favoured low-energy binding site by a fast, reversible non-covalent association of the molecules in the minor groove through interactions including hydrogen bonding, van der Waals and electrostatic interactions[90]. If these non-covalent intermolecular interactions are weak, the PBD presumably dissociates and re-associates at another site, with the process repeating itself until a suitable low-energy triplet with the C2-NH_2 of the central guanine aligned for nucleophilic attack at the PBD C11-position is found. In the second step, the PBD forms a covalent bond with the guanine base, and the molecule is then locked into position. Whereas the initial non-covalent association is a fast reaction, the covalent attachment step is much slower and can take up to 24 hours to complete[89b] (**Figure 1.21**).

Once covalently bonded to DNA, PBD monomers and dimers have been shown to mediate a number of biological effects in cells including DNA strand breakage[94], inhibition of endonucleases[100] and RNA polymerase[89b, 101], and transcription factor inhibition. For example, some PBDs can prevent the transcription of genes crucial for cell survival such as NF-Y[102] and NF-κB[90].

1.4.1 Structure Activity Relationships (SARs)

A large body of Structure Activity Relationship (SAR) data has been generated based on both the known natural products and a large number of synthetic compounds that have become available. A summary of the key SAR features of the PBD monomers (**Figure 1.22A**) and dimers (**Figure 1.22B**) is provided below.

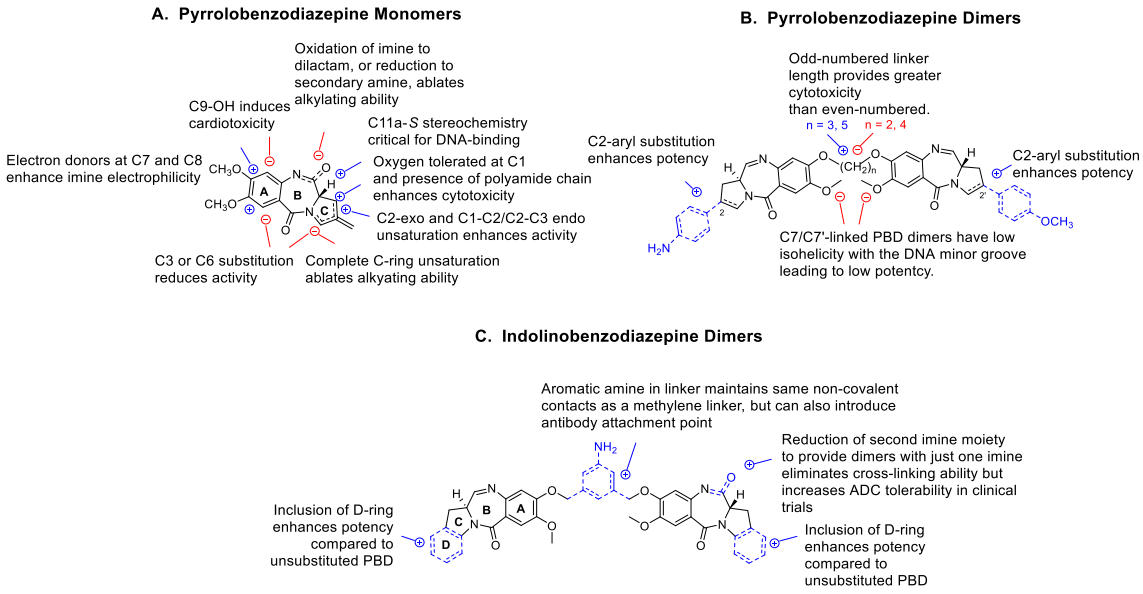


Figure 1.22: Summary of the Structure Activity Relationships (SARs) for the PBD monomers (**A**), the PBD dimers (**B**), and the indolinobenzodiazepine analogues (**C**). Modifications which enhance activity are shown in blue, and those which reduce activity are in red.

N10-C11 imine (or equivalent)

The presence of an imine, carbinolamine or carbinolamine methyl ether group at the N10-C11 position is essential for the covalent interaction of a PBD molecule with DNA. Early studies by Kohn[103] and Hurley[104] demonstrated that the PBDs covalently bond to guanine-containing sequences, and also established that they do not interact through an intercalation mechanism[105] (**Figure 1.21**).

Crucially, replacement of the N10-C11 imine with amide or secondary amine functionalities (**Figure 1.23A, B**) results in a dramatic fall in DNA-binding ability[106] and cytotoxicity. For example, it has been demonstrated that DNA-binding affinity upon replacing the two N10-C11 imines of the PBD dimer SJG-136 (**6a**) with amide functionalities is lost[94]. Similarly, didehydroanthramycin[107] (**Figure 1.23C**) which has extended N10–C11/C11a–C1/C2–C3 conjugation is biologically inactive, as the N10-C11 imine is too stable to react with a guanine base. Although PBD dilactams (*e.g.*, **4a**) cannot bind covalently to DNA, they are still isohelical with the DNA minor groove due to their chiral C11a-position, and so can possess weak DNA-binding properties through

non-covalent hydrogen bonding and other interactions such as van der Waals. For example, a series of dilactams synthesised by Lown and co-workers[108] provided a ΔT_m of up to approximately 3 °C with calf thymus DNA, although this compares unfavourably with a simple PBD monomer such as anthramycin which gives a ΔT_m of 13.1 °C under the same conditions, and PBD dimers produce significantly higher ΔT_m values due to their length (and hence greater number of non-covalent contacts in the minor groove) and cross-linking ability[100]. Libraries of PBD dilactams synthesised by Thurston and co-workers have shown similar ΔT_m value ranges (*e.g.*, up to approximately 2.4 °C for C2-aryl substituted dilactams)[106a].

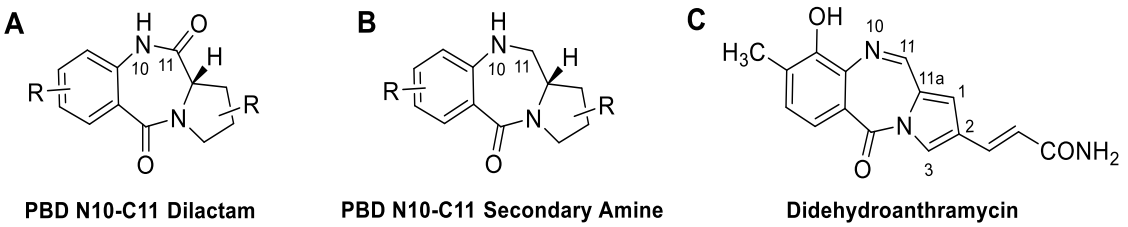


Figure 1.23: Structures of PBD analogues in which the electrophilic N10-C11 imine moiety has been replaced with an amide (*i.e.*, dilactam, **A**) or a secondary amine (**B**), or in which the imine has been rendered non-electrophilic by conjugation with double bonds in the C-ring to create a fully aromatised system (*i.e.*, didehydroanthramycin, **C**). All molecules of types **A**, **B** and **C** cannot bond covalently with DNA, and so have negligible cytotoxicity. Some examples (*e.g.*, the dilactams) can still bind to DNA, albeit weakly, through non-covalent interactions because they still have the appropriate shape to fit into the DNA minor groove[106a].

C11a/C11S stereochemistry

It has been shown through NMR[109] and X-ray crystallography[110] studies that all PBDs exist in a single stereochemical form (*S*) at the C11a-position which provides a right-handed twist that allows the molecule to follow the curvature of the minor groove within the right-handed B-DNA helix. Molecules with the (*R*)-stereochemistry at C11a are not isohelical with the DNA minor groove, causing them to sterically interact with the groove wall[110-111]. Examples of PBDs with (*R*)-stereochemistry at C11 have been synthesised and, as predicted, do not possess any significant DNA-binding affinity or cytotoxicity[112].

Conflicting data exists in the literature as to the stereochemistry at the C11-position of the PBD/guanine adduct. Although the anthramycin crystal structure[110] and a number of NMR studies of both PBD monomers[109, 111] and PBD dimers[113] have shown that the adduct formed is in the C11-*S* configuration, molecular mechanics calculations [114] and an NMR structure[115] have suggested that in some situations, both *R* and *S* configurations are equally preferred. It is now thought that PBD monomers can form either *R* or *S* adducts at C11, and this is likely to relate to the substitution pattern of a particular PBD (*e.g.*, C8-substitution may be partially influential) and the sequence to which it is covalently bound[114c].

Functionalization of the A-ring

The number and pattern of substituents in the A-ring of the PBD skeleton can affect the DNA-binding ability and cytotoxicity. The presence of electron-donating substituents, particularly at the C7- and C8-positions, can enhance DNA interactivity by increasing the electrophilicity and alkylating potential of the N10-C11 imine moiety[116]. Furthermore, extensive molecular modelling studies have shown that substituents at the C8-position position themselves along the floor of the minor groove which can lead to significant additional stabilisation of the adduct[90]. Thus, polyheterocyclic fragments (*e.g.*, pyrroles, furans, imidazoles, thiophenes, indoles, benzofurans and biaryl units) have been added to this position to create PBD C8-conjugates (*e.g.*, GWL-78, KMR-28-39, **Figure 1.20**). Similarly, two PBD units have been joined through their C8-positions to create C8/C8'-linked PBD dimers[94, 113a, 117], and this approach led to the design and synthesis of the C8/C8'-linked PBD dimer SJG-136[118] (**Figure 1.20**). The C7-position has also been used as a potential linker point for PBD dimers[93a], but molecules joined in this way have significantly less DNA cross-linking ability and cytotoxicity, as the two PBD units are not properly aligned for alkylation of guanine bases[93b].

The presence of a hydroxyl group at C9 renders the molecule cardiotoxic[119], which is thought to be due to free radical production in cardiac tissue[120]. Other studies on PBDs such as anthramycin[110] have shown that a C9-OH has the potential to hydrogen bond to bases within the minor groove of DNA, thus increasing binding affinity and sequence selectivity[110]. The presence of a larger group (*e.g.*, a methoxy group) at C9 renders the PBD biologically inactive, as the steric hindrance between this substituent and functional

groups on the minor groove floor inhibits alkylation as can be demonstrated through molecular dynamics simulations (Jackson and Thurston, *unpublished data*).

C-ring unsaturation

The degree of saturation/unsaturation within the C-ring can also influence DNA-binding ability and cytotoxicity. In general, PBD monomers with C2-*exo* or C2-*endo* unsaturation (*e.g.*, anthramycin[104], tomaymycin[121] and sibiromycin[122]) have enhanced DNA-binding ability relative to PBD monomers with fully saturated C-rings (*e.g.*, neothramycin A[123], chicamycin[124] and DC-81[125]). This feature has been adopted in the rational design of PBD dimers (**Figure 1.22B**), where the introduction of *exo*-unsaturation at the C2/C2'-positions of the PBD dimer DSB-120[113a] (**5a**, **Figure 1.20**) led to SJG-136 (**6a**, **Figure 1.20**). DSB-120 has poor activity *in vivo*, attributed partly to its high reactivity with cellular thiol-containing molecules such as glutathione[126]. However, the introduction of C2-*exo* unsaturation as in SJG-136 led to a decrease in electrophilicity of the N10-C11 imines and an overall increase in DNA-binding affinity and cellular cytotoxicity, thought to be due to a lower reactivity toward cellular nucleophiles with more of the agent reaching its target DNA. In addition, modeling studies have shown that C2/C2' unsaturation causes a flattening of the C-ring which may lead to superior van der Waals contacts within the minor groove[94] and enhanced DNA-binding affinity[114a]. Interestingly, complete unsaturation of the C-ring significantly reduces the electrophilicity of the N10-C11 imine by producing a fully aromatic system across N10-C11/C11a-C1/C2-C3 which ablates DNA-binding ability and cytotoxicity completely (*e.g.*, didehydroanthramycin[107], **Figure 1.22C**).

C2-C3 substitution

In a similar manner to C9-substitution, a substituent at C3 can affect the DNA-binding of a PBD in the minor groove. For example, methylation or butylation of the C3-hydroxyl substituent of neothramycin A blocks interaction in the DNA minor groove[127]. However, extended C2-substituents (*e.g.*, the polyacrylamide side chain of anthramycin[110]) significantly enhances DNA interactivity because they locate along the minor groove and stabilise the adduct through van der Waals interactions and sequence-specific hydrogen bonds with side chain and functional groups on the minor groove floor. For example,

early studies on anthramycin suggested that removal of the C2-acrylamide tail can significantly reduce its DNA interactivity[128]. This feature has been used in the design of novel C2-substituted PBD monomers with enhanced DNA-binding and cytotoxicity[106a]. For example, some C2-substitutents such as -phenyl-CH_3 moieties can lead to cytotoxicities in the low nanomolar range[106a]. Studies have also shown that replacement of the C1-atom itself with an oxygen provides analogues (*e.g.*, oxazolbenzodiazepines) with similar activity to the parent PBDs[89a, 127, 129].

Many of the modifications outlined above have been incorporated into the PBD dimers that are now being used as chemical payloads for Antibody-Drug Conjugates (ADCs) (**Figure 1.22B** and **C**). For example, the ADC SGN-CD33A contains a C2-C3 *endo*-unsaturated C2-aryl-substituted PBD with a C8/C8'-propylenedioxy linker[99]. The A-ring contains oxygen substituents in both the C7- and C8-positions which enhance its interactivity with DNA, and the C3-, C6- and C9-positions remain unsubstituted, thereby minimising any steric hindrance within the DNA minor groove. In addition, researchers at the company ImmunoGen Inc. have added a fourth D-ring to the C-ring of the PBD skeleton to produce indolinobenzodiazepines (**Figure 1.22C**)[130]. They have also introduced an aromatic ring into the centre of the C8/C8'-linker which does not affect DNA-binding affinity but provides an antibody attachment point (**Figure 1.22C**)[131].

1.4.2 Evolution of the PBD dimers

The first PBD dimers were synthesised by Suggs and co-workers in 1988[93b] with the two PBD units joined through linkers attached at the C7/C7'-positions of the aromatic A-rings (**4a**-**c**, **Figure 1.20**). Although modest DNA cross-linking activity was reported, no cytotoxicity data was provided. Based on a rational approach to the way in which a PBD dimer might best fit into the DNA minor groove, in 1992 Thurston and co-workers[132] reported the PBD dimer DSB-120 (**5a**, **Figure 1.20**) which consists of two unsubstituted PBD units coupled through their C8/C8'-positions *via* a 1,3-dioxypropyl linker. This molecule, which was designed to more closely follow the curvature of the DNA minor groove, had very high DNA-stabilising (*i.e.*, ΔT_m = > 15.1 °C) and interstrand cross-linking (*i.e.*, CL = > 90%) properties, and significant cellular cytotoxicity[132]. SAR studies[133] demonstrated that changing the central methylene linker from 3 (**5a**, **Figure**

1.20) to 4, 5 or 6 methylene units (**5b-d**, **Figure 1.20**) still allowed DNA-binding but significant differences were observed. Increasing the linker length to n = 5 (**5c**, **Figure 1.20**) afforded a molecule with similar cross-linking efficacy to DSB-120 (n = 3) but with approximately half of the DNA-stabilising ability, while the cellular cytotoxicity was significantly enhanced. Conversely, increasing the linker length to n = 4 or 6 afforded a reduction in all three parameters, and led to the conclusion that linkers with an odd number of methylenes are preferred for maximum activity. This was later explained through molecular modeling studies[133], as only dimers with odd numbers of methylenes in their linkers allow the two PBD units to arrange themselves in the correct orientations for their N10-C11 imine moieties to alkylate guanine C2-NH_2 groups. Although DSB-120 had potent activity in biophysical and cellular assays, it proved to be inactive in human tumour xenograft studies in mouse models. This was thought to be due to the N10-C11 imine moieties being too electrophilic due to the fully saturated state of the C-rings, with the molecule reacting with extra- and intra-cellular thiol-containing nucleophiles before it could reach its DNA target[134]. To reduce the electrophilicity of the N10-C11 imine moieties and increase the chances of the molecule reaching the target DNA in the nucleus of cells, methylene substituents were added at the C2/C2'-positions to afford SJG-136 (**6a**, **Figure 1.20**)[90, 96, 135]. This analogue was equally potent in biophysical and cellular assays but was also active in *in vivo* assays. This was attributed to the flattening of the C-rings due to the presence of the C2/C2'-substituents which modifies the electrophilicity of the N10-C11 imines through an, as yet, unidentified mechanism, but which also provides a better fit of the molecule in the minor groove.

Due to the presence of two electrophilic imine moieties, SJG-136 (**6a**, **Figure 1.20**) can form interstrand and intrastrand cross-links in addition to mono-alkylated adducts[94] (**Figure 1.24**). The type of adducts that form appears to depend on the precise DNA sequence and the distance between reacting guanines[96, 136]. Studies have shown that SJG-136 (**6a**, **Figure 1.20**) has a preference for forming Pu-GATC-Py interstrand cross-linked adducts[96, 137].

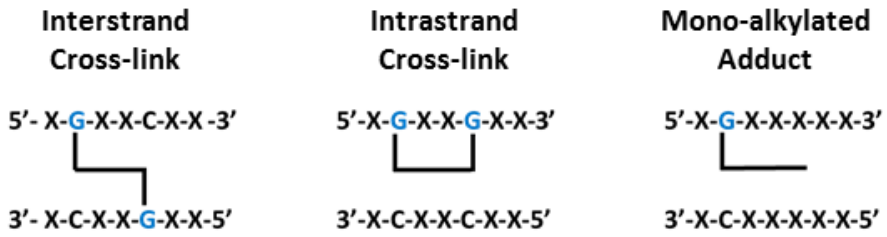


Figure 1.24: Diagram of the various possible adduct types that can be formed between SJG-136 (**6a**, **Figure 1.20**) and DNA (*i.e.*, interstrand and intrastrand cross-links, and mono-adducts).

Examples also exist in the literature of PBD dimers linked through the C8/C2'-[95a] or C2/C2'- positions rather than the C8-C8'-positions (**Figure 1.25**), but these molecules are at least 10^3-fold less cytotoxic than the most potent C8-linked dimers. For example, Thurston and co-workers[95a] synthesised a C8/C2'-linked PBD dimer (**Figure 1.25A**) containing C2'/C3'-*endo* unsaturation at the C2-linkage, which had an IC_{50} of > 25 μM in A2780 cells (human ovarian carcinoma cell line). This is less cytotoxic than the core PBD monomer DC-81, suggesting that the second PBD unit was impeding rather than enhancing DNA-interaction. Kamal and co-workers[138] synthesised a similar C8/C2'-linked dimer consisting of a C8'-phenyl-PBD linked *via* its C2'-position to the C8 of a second PBD unit (**Figure 1.25B**). This dimer also had cytotoxicity in the micromolar range (*e.g.*, 2.6 μM in MCF-7 cell lines, human breast cancer), and in DNA thermal denaturation studies provided a ΔT_m of 5.7 °C after incubation with calf thymus DNA for 18 hours, which compares unfavourably with the ΔT_m for the C8/C8'-linked dimer DSB-120 under similar conditions (15.1 °C), again suggesting a poor fit in the DNA minor groove. In the same study, Kamal and co-workers synthesised an analogue which had one of its N10-C11 imine functionalities converted to an amide (**Figure 1.25C**). Interestingly, this molecule produced a level of cytotoxicity approximately 40-fold less than the *bis*-imine parent molecule (**Figure 1.25B**) suggesting that, although the molecule could no longer cross-link DNA, some form of DNA interaction (presumably mono-alkylation) could still occur. Molecular dynamics studies have since confirmed that the 3-dimensional shape of C8/C2'-linked PBD dimers is not conducive to a good fit within the DNA minor groove[139].

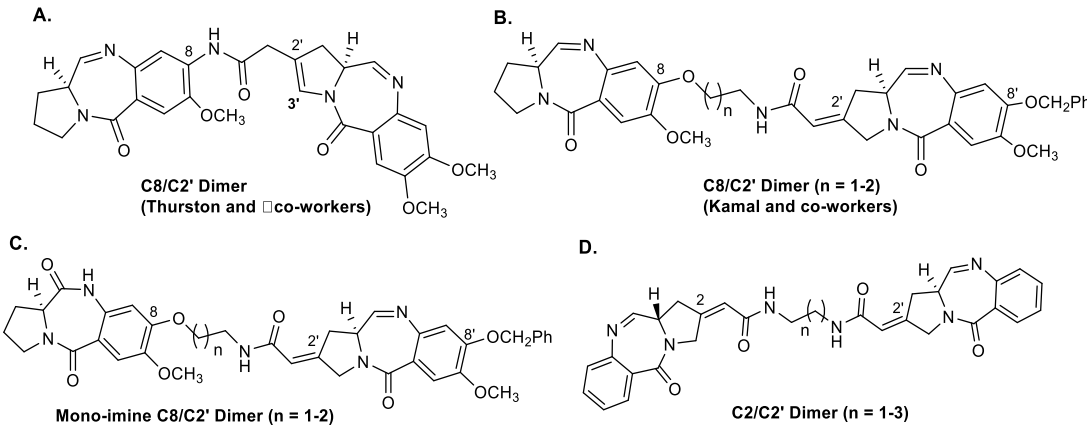


Figure 1.25: Examples of C8/C2'- and C2/C2'-linked PBD dimers synthesised by the Thurston[95a], Kamal[95b] and Lown[95b] groups. Dimer **1.25C** differs from the others in that it possesses only one electrophilic N10-C11 imine moiety, and so cannot cross-link DNA.

A library of C2/C2'-linked dimers has also been synthesised by Lown and co-workers[95b] (**Figure 1.25D**). Cytotoxicity studies indicated that one member of the series (where n = 3) had an average LC_{50} of approximately 50 μM across 60 cancer cell lines. Molecular dynamics simulations have since confirmed that PBD dimers linked in a C2/C2' manner do not have the appropriate 3-dimensional for effective interaction in the DNA minor groove[139].

1.4.3 The PBD dimer SJG-136

The best known PBD dimer, SJG-136 (**Figure 1.26**) (also known as NSC 694501 or BN2629), was synthesised in the 1990s and is currently being investigated in Phase II clinical trials in ovarian cancer and leukaemia patients[98b, 140]. More recently, PBD dimer analogues are being attached to tumour-targeting antibodies to create Antibody-Drug Conjugates (ADCs), a number of which are now in early clinical trials, with many others in pre-clinical development.

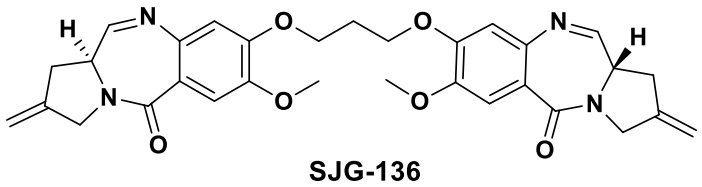


Figure 1.26: Chemical structure of the PBD dimer SJG-136.

SJG-136 is composed of two PBD monomer units connected through a C8-O-$(CH_2)_3$-O-C8' linker[90, 96, 135]. Both of the PBD monomers contain an electrophilic imine moiety at the N10-C11 position which allows SJG-136 to form interstrand crosslinks in addition to intrastrand, and mono-alkylated adducts[94]. This adduct type preference is depending on the DNA sequence and the distance between reacting guanines. Studies with SJG-136 have shown that the agent has a preference for the Pu-GATC-Py interstrand cross-linked adduct[96, 137].

An increasing body of evidence suggests that PBD monomers exert their pharmacological effect through transcription factor (TF) inhibition. The PBD monomer GWL-78 has been shown to inhibit the TF NF-Y[102] and KMR-28-39 has recently been reported to interact with one or more consensus DNA sites of the TF NF-κB[90]. These findings led to the proposal that TF inhibition may contribute to the overall anti-tumour activity of SJG-136 within the biological system and also may explain the observed differences in its activity in various human tumour cell lines.

1.5 Biophysical methods used to study DNA-binding agents

Newly discovered DNA-binding agents are initially evaluated in simple *in vitro* experiments, including UV spectroscopy[141], fluorescence[141], Raman spectroscopy, circular dichroism[142] and DNA thermal denaturation studies[143] in order to determine their ability to bind to DNA. However, these techniques are useful for the initial assessment of the DNA-binding capability of those ligands and do not provide any additional information regarding the sequence specificity.

The DNase footprinting assay is the most widely used technique to identify the sequence selectivity of novel anti-tumour agents[144]. In this assay, a DNA fragment of known sequence is labelled with ^{32}P and incubated with DNaseI in the presence of the ligand. By binding to specific sequences on the DNA, the ligand protects those from enzymatic digestion by DNaseI resulting in a differential cleavage pattern which is subsequently visualised using agarose gel electrophoresis. This technique has been applied to study the sequence preference of netropsin[145] and chlorambucil[146].

Furthermore, the single cell electrophoresis assay (COMET) has become established as an uncomplicated and highly sensitive technique for the measurement of DNA strand breaks at the level of the individual eukaryotic cell[147]. Cells are suspended in agarose and casted on a microscope slide and lysed with detergents and high salt. During this process, undamaged DNA forms highly organised structures with the matrix proteins in the nucleus, whereas damaged DNA loses this property. During electrophoresis at high pH, the undamaged DNA is not able to migrate through the agarose gel towards the positively charged anode due to the supercoiled loop structure. DNA strand breaks, however, are capable of leaving the cavity and migrating towards the anode as they are composed of smaller fragments. Therefore, the amount of DNA that migrated during electrophoresis is representative for the amount of DNA damage in the cell. Image analysis compares the overall intensity of the fluorescence signal for the sample to the signal of the migrated DNA.

1.6 The evaluation of SJG-136

Extensive pre-clinical and clinical evaluation has been carried out on SJG-136. The most relevant methods and results are described in the following section.

1.6.1 Biophysical evaluation of SJG-136

The DNA binding affinity of SJG-136 was investigated in DNA thermal denaturation studies using calf thymus DNA at SJG-136/DNA ratios of 1:100, 1:50 and 1:5[94]. The ΔT_m values obtained (**Table 1**) demonstrated an increase in the DNA melting temperature for each ratio after 0, 4 and 18 hour of incubation at 37 °C, clearly indicating the ability of SJG-136 to stabilise double-stranded calf thymus DNA. The data show a significant elevation of helix melting temperature by 33.6 °C after 18 h incubation at a 1:5 ratio (SJG-136/DNA), with a significant stabilisation effect already evident at t = 0, suggesting a strong kinetic component to the interaction. After 72 h incubation at a 1:5 ratio, the ΔT_m increased still further to a value of 34.4 °C, providing a plateau saturation level for covalent DNA interaction. It is noteworthy that the ΔT_m values obtained were significantly higher than for the PBD monomers tomaymycin and sibiromycin, with sibiromycin providing the highest ΔT_m value at 18 hours for all of the known PBD monomers[122].

Compound	**[PBD]:[DNA] molar ratio**	**Induced ΔT_m (°C) after incubation at 37 °C for:**		
		0 h	**4 h**	**18 h**
SJG-136	1:100	7.1	8.0	9.1
SJG-136	1:50	11.3	12.3	15.0
SJG-136	1:5	25.7	31.9	33.6
Tomaymycin	1:5	0.97	2.38	2.56
Sibiromycin	1:5	-	-	16.3

Table 1: Thermal denaturation data for the interaction of SJG-136 with calf thymus DNA after incubation at different ratios for 0, 4 and 18 hours at 37 °C[94, 100].

The extent of interstrand cross-link formation produced by SJG-136 was investigated using a gel electrophoresis assay developed by Hartley and co-workers[94, 148]. It was evaluated at concentrations between 0.001 to 10 μM, and demonstrated highly efficient DNA cross-linking (**Figure 1.27a**). After 2 h incubation of DNA with SJG-136, cross-links were visible starting from a concentration as low as 0.01 μM, with single-stranded DNA disappearing completely at 0.3 μM. The degree of cross-linking was quantitated using laser densitometry as shown in **Figure 1.27b**. Overall, these results demonstrated the very high potency of SJG-136 as a DNA interstrand cross-linking agent.

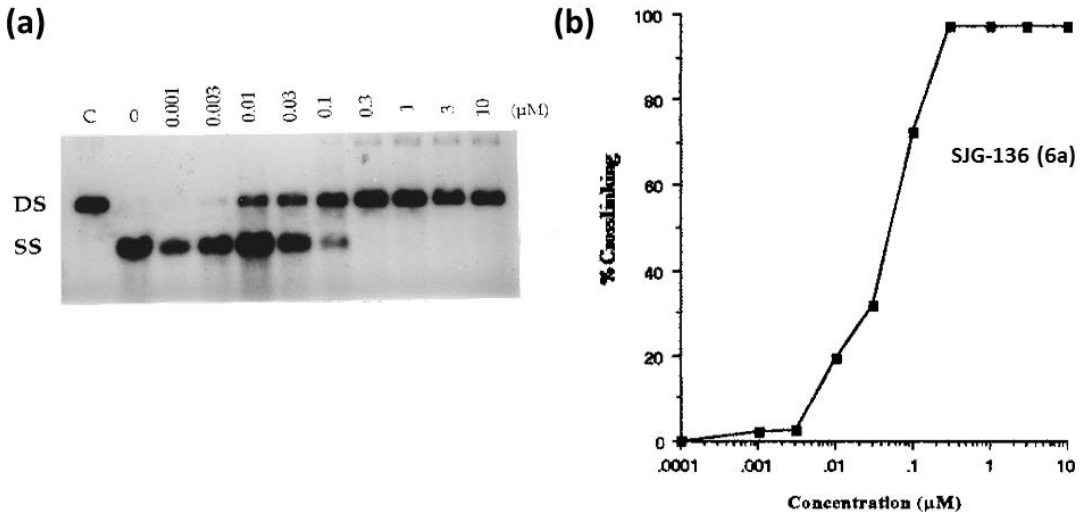


Figure 1.27: (**a**) Autoradiograph of a neutral agarose gel showing DNA interstrand cross-linking of linear ^{32}P-end-labeled pBR322 DNA by SJG-136 following a 2 h incubation at 37 °C. The lanes are: C, double-stranded DNA control; 0, single-stranded DNA control; and 0.001, 0.003, 0.01, 0.03, 0.1, 0.3, 1.0, 3.0, 10.0 μM SJG-136. DS and SS are double- and single-stranded DNA, respectively; (**b**) Quantification of the gel in **Figure 1.27a** to show the concentration dependence of DNA cross-linking for SJG-136 in linear ^{32}P-end-labeled pBR322 DNA[94]. "Reprinted with permission from Gregson, S. J., *et al* (2001) Design, synthesis, and evaluation of a novel pyrrolobenzodiazepine DNA-interactive agent with highly efficient cross-linking ability and potent cytotoxicity, *Journal of Medicinal Chemistry 44*, 737-748. Copyright © 2001 American Chemical Society."

Molecular modeling studies of cross-link formation were carried out to support the observations made in the thermal denaturation and DNA cross-linking studies[94]. Two 12-mer DNA duplexes, 5'-CGCA***GATC***TGCG-3' and 5'-CGC***TCTAG***AGCG-3' (G-G cross-

link sequences spanned underlined; alkylated guanine bases in bold), were designed for energy calculations. Each sequence offered optimal binding sites (*i.e.*, 5'-A**G**A) on either strand for alkylation by a PBD but in either the 3'- or 5'-direction. The results suggested that either no (or insignificant) disruption to the DNA secondary structure occurred upon adduct formation and cross-linking, and that no induced distortion of base pairs within the helical structure occurred[94]. Furthermore, energy-minimised models demonstrated excellent accommodation of SJG-136 within the DNA minor groove, with no significant part of the molecule exposed beyond the periphery of the duplex. This feature may explain the observed resistance of SJG-136 adducts to DNA repair which is usually initiated after repair enzymes recognise distortion or helical perturbation of the DNA helix as occurs with the majority of DNA alkylating agents. The calculated binding energies also supported the observation that covalent interaction between SJG-136 and DNA stabilises the DNA structure to a greater extent than the related PBD monomers (*i.e.*, the energy of binding values were twice as low as for the monomers). Moreover, the flexible C8/C8'-linker which connects the PBD monomeric units in SJG-136 was observed to enhance overall binding in the DNA minor groove through van der Waals interactions, with hydrogen bonds observed to form between the N10-protons of the PBD units and the N3-atoms of 3'-adenine bases adjacent to the covalently-modified guanines on either strand. These computational studies allowed models to be created for the interaction of SJG-136 with DNA as shown in **Figure 1.28**. Overall, these modelling results were consistent with the high ΔT_m values observed for SJG-136 compared to the PBD monomers (**Table 1**).

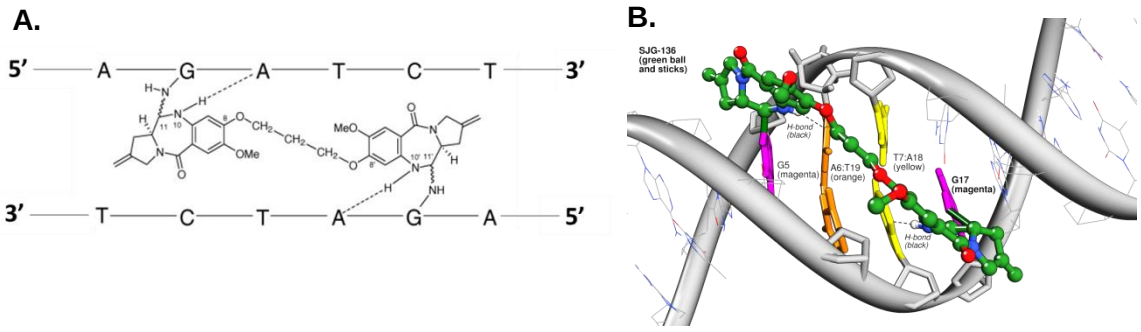


Figure 1.28: **A**. Schematic diagram of the PBD dimer SJG-136 covalently binding to, and interstrand cross-linking with, the DNA sequence 5'-AGATCT-3'. Its sequence-selectivity is due in part to (a) covalent bonds forming between the C11/C11' positions of the PBD units and the

exocyclic N2 groups of guanines on opposite strands (shown as ----), and (b) hydrogen bonds (shown as - - -) formed between the N10/N10' protons of the PBD units and the ring nitrogen N3 acceptors of the 3'-adenines adjacent to the covalently modified guanines[149]; **B**. Molecular model of SJG-136 forming an interstrand cross-link within the sequence 5'-TATAGATC(G)TATA-3' (covalently modified bases underlined, and bases on the opposing strand in brackets) between the guanines on each strand (both magenta).

The sequence-selectivity of SJG-136 toward naked DNA was initially studied using a traditional DNA footprinting method based on a 262-bp duplex DNA sequence from the MS2 plasmid[150]. This DNA fragment contained the preferred cross-linking site for SJG-136 (*i.e.*, 5'-Pu-GATC-Py-3') in addition to less-favoured sites (*i.e.*, 5'-GXXC-3'; X = any base)[94]. The sequence included two preferred binding sites at positions 5'-^{48}GGATCC^{53}-3' and 5'-^{208}GGATCC^{213}-3', together with 17 other less-favoured potential binding sites containing the sequence 5'-GXXC-3'. As anticipated, the two predicted most-energetically-favoured binding sites were 5'-^{48}GGATCC^{53}-3' and 5'-^{208}GGATCC^{213}-3', where the molecule was observed to bind with high affinity. It was also observed to bind to a number of lower affinity sites such as 5'-GXXC-3', where XX was TA, GC, CT, TT, GG, or GT. This result was surprising at the time of publication, given that a guanine residue should not be well-tolerated between the cross-linked guanines due to the bulky nature of the C2-exocyclic amino group which points into the minor groove. However, based on more recent studies that have demonstrated the formation of mono-adducts and intrastrand cross-links in addition to interstrand cross-links[96], this observation can now be explained by the formation of these other adduct types. Interestingly, the gels also showed high molecular weight material appearing from 1 μM SJG-136 concentration and above at the top of the gel, which intensified with higher concentrations. This material was assumed to be duplex DNA that could not denature to single stranded DNA either before and/or during denaturing gel electrophoresis due to the interstrand cross-links formed by SJG-136.

Apart from the effect of DNA interstrand cross-linking on DNA replication, it is also possible that the *in vitro* and *in vivo* biological activity[151] of SJG-136 may be due to its sequence-selective interaction with DNA binding sites corresponding to key sequences in genes involved with the promotion of tumour growth. Therefore, an *in vitro* transcription stop assay was carried out on a 282-bp DNA sequence to investigate the

effect of SGJ-136 on T7 RNA polymerase and transcription termination[101, 150]. Significant transcriptional stop sites (T-stops) were observed which appeared to be mainly associated with 5'-GXXC-3' sequences. Although this assay cannot distinguish between mono-alkylated and cross-linked adducts, at the time the T-stops produced were assumed to be mainly associated with interstrand cross-links, as high affinity adducts would be required to produce significant T-stops[101]. In accord with the DNA footprinting experiments, the most favourable sequence 5'-^{48}GGATCC^{53}-3' gave the strongest T-stop, suggesting that this sequence might be the principal cross-linking site for SJG-136. The kinetic behaviour of SJG-136 was also studied by time-dependent experiments, and the results showed that incubation with DNA for 15 minutes was sufficient to inhibit transcription with negligible changes up to 1 hour incubation. Intriguingly, the data also revealed that, once transcription was initiated, addition of SJG-136 had no effects on transcription inhibition, suggesting that adducts needed to be fully formed before they could block transcription. In addition, temperature-dependent experiments showed that the activation energy for covalent drug-DNA interaction must be very low at 5'-Pu-GATC-Py-3' sites, since footprints were clearly detectable on the gels at the same site at 4 °C. In contrast, the activation energy for the less preferred 5'-^{39}GGTACC^{44}-3' site must be higher, as the strongest footprint was observed at only 37 °C.

A significant number of studies were also carried out in tumour cell lines (both human and canine) using the Single Cell Electrophoresis (COMET) and/or γ-H2AX Foci assays to correlate the extent of DNA cross-linking with cytotoxicity[149]. SJG-136 was initially evaluated in K562 (leukaemia), U266B1 (myeloma), U937 (lymphoma), PC-3 (prostate), OVCAR-5 (ovarian), MCF-7 (breast), CaCo2 (colon) and HCT-15 (colon) cell lines using the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazol (MTT) growth inhibition assay, and an average GI_{50} (concentration for 50% maximal inhibition of cell proliferation) of 212 pM was established (range: 2.1 pM in A2780 to 2.3 nM in HCT-15)[149]. The COMET and/or γ-H2AX Foci assays were then used to detect DNA interstrand cross-links in some of these cell lines following a one hour exposure to SJG-136. Cross-linking was detectable at very low SJG-136 concentrations (*i.e.*, 0.5 nM in the K562 cell line) with cross-links persisting over a 48 hours period.

The γ-H2AX assay[152] is more sensitive than the COMET assay[153] in its ability to detect DNA damage caused by DNA cross-linking agents. γ-H2AX foci formation is rapidly

induced at DNA double-strand breaks and sites of replication fork collapse that are often induced by cross-linking agents[154]. They accumulate at the sites of DNA damage and serve to recruit DNA repair and cell cycle checkpoint enzymes in order to initiate repair of the damage. They can be visualised using immunostaining techniques, and the number of γ-H2AX foci present is proportionate to the DNA damage. Therefore, the γ-H2AX assay has become an important and useful biomarker assay for measuring the DNA damage caused by DNA cross-linking agents such as SJG-136[155]. The assay was initially used to establish whether it could detect a correlation between the rate of interstrand cross-link formation and the appearance of γ-H2AX foci in HCT-116 cells[156]. The cells were exposed to either 50 nmol/L SJG-136 for 1 hour, or 1 nmol/L for 24 hours, and significant foci formation was detected for both conditions after 4 hours or 8 hours post treatment, respectively. Despite these exposure differences, the levels of γ-H2AX foci formed in both cases were comparable after 24 hours, demonstrating that interstrand cross-links (ICLs) are formed earlier (*i.e.*, maximum reached after 1 hour SJG-136 treatment) than γ-H2AX foci are detected (*i.e.*, visible after 4 hours exposure to the agent) when the cells were exposed for 1 hour to a higher concentration of SJG-136. This delay was even more pronounced in cells treated with a lower concentration of SJG-136 for 24 hours. In this case, H2AX foci were visible at 8 hours post incubation whereas cross-link formation could be observed 1 hour after a one hour exposure, with maximum ICL formation after 24 hours. These results were consistent with the need for DNA adducts to form prior to γ-H2AX foci appearing.

A reversed phase HPLC/MS methodology has also been developed to study the sequence selectivity of SJG-136[157]. This technique requires only small quantities of short inexpensive oligonucleotides, and does not require the use of radioactivity compared to traditional methods, such as DNA footprinting. Furthermore, the adducts formed can be rapidly visualised (*i.e.*, < 30 min), and so both thermodynamic and kinetic data can be obtained, along with sequence selectivity information if different oligonucleotides are used. Initial studies of SJG-136 using this methodology involved self-complementary 12-mer oligonucleotides containing Pu-GATC-Py (*Seq-1*), Pu-GTAC-Py (*Seq-2*), Py-GATC-Pu (*Seq-3*) and Pu-IATC-Py (*Seq-4*) sequences (I = Inosine). This assay was also used to investigate the extent of cross-link formation at different molar ratios after 24 hours. The data obtained clearly showed that optimal cross-linking (*i.e.*, ~100%) occurred at a 4:1 (SJG-136/DNA) ratio[157]. The fastest reaction was observed with *Seq-1*, followed

by *Seq-2* and *Seq-3*, with *Seq-4* not reacting. These results confirmed that SJG-136 preferred the sequence Pu-GATC-Py (*Seq-1*), consistent with the previous gel-electrophoresis studies[88b, 94]. *Seq-4*, which had its guanines replaced with non-nucleophilic inosines, failed to react as anticipated, thereby acting as a negative control.

The biological activity of SJG-136 was thought to be entirely due to the Pu-GATC-Py interstrand cross-links formed in the DNA minor groove[94, 150, 158]. However, based on the results of further HPLC/MS studies, it became clear that SJG-136 is capable of forming longer interstrand cross-links at Pu-GAATC-Py sequences, intrastrand cross-links at both shorter (*e.g.*, Pu-GATG-Py) and longer (*e.g.*, Pu-GAATG-Py) sequences, and mono-alkylated adducts at guanine residues where cross-linking is not possible because of the distance of the next potentially reacting guanine[96]. Furthermore, a direct competition experiment was then carried out between the interstrand Pu-GATC-Py and Pu-GAATC-Py sequences, and the intrastrand Pu-GATG-Py and Pu-GAATG-Py sequences. The results indicated a rank order of reactivity of Pu-GAATG-Py > Pu-GATC-Py »Pu-GATG-Py > Pu-GAATC-Py.

Molecular dynamics simulations were undertaken to gain insight into these observations[96]. It was clear from energy minimised models that the two electrophilic N10-C11 imine moieties of SJG-136 were well-positioned to react covalently with the guanines in the Pu-GATC-Py (*i.e.*, favoured interstrand cross-link) and Pu-GATG-Py (*i.e.*, intrastrand cross-link) sequences with little distortion of the DNA helix. However, for the extended Pu-GAATC-Py and Pu-GAATG-Py sequences, the model predicted that bis-alkylation could only occur at the expense of some distortion of the DNA at the points of covalent attachment, and with SJG-136 adopting a slightly lower position in the DNA minor groove in order to span the necessary distance between reactive guanines.

1.6.2 Biological evaluation of SJG-136

In the late 1990s, SJG-136 was evaluated in the National Cancer Institute's (NCI's) 60-cell line panel in which it showed considerable promise in terms of potency and selectivity[82, 90]. This screen, which comprised sub-panels of leukaemia, melanoma, lung, colon, kidney, ovary and central nervous system tumour cells was, in its original form, capable of evaluating more than 10,000 new chemical agents per year, and was used as

the basis to select both compounds and cell lines for follow-up Hollow Fibre and *in vivo* xenograft studies. The results indicated a LC_{50} (*i.e.*, median lethal concentration) of 7.4 nmol/L with a range of 0.14 to 320 nmol/L. In particular, the significant differences between the LC_{50}, TGI (*i.e.*, tumour growth inhibition) and GI_{50} (*i.e.*, concentration for 50% maximal inhibition of cell proliferation) values across the different cell lines suggested that SJG-136 had selective cytotoxicity toward certain panels of cell lines and individual cell lines. For example, the leukaemia and melanoma cell panels were particularly sensitive. Furthermore, it was possible to compare the overall cytotoxicity profile of SJG-136 with those of approximately 60,000 compounds in the NCI's database using a molecular target analysis program known as COMPARE[159]. Although this analysis indicated an activity pattern for SJG-136 similar to some known DNA-binding agents (*e.g.*, melphalan, cyclophosphamide, chlorambucil), its overall profile was not a perfect match with cluster patterns associated with any specific chemotherapeutic agents, thus reflecting its unique mechanism of action.

Based on the results from the NCI's 60-cell line screen, SJG-136 was next evaluated in the Hollow Fibre assay[158] to guide the selection of tumour cell types for use in xenograft studies. This assay involves cultivation of sensitive tumour cells in porous hollow fibre tubes which were then implanted into the intraperitoneal (i.p.) and subcutaneous (s.c.) compartments of host immunocompromised mice. These cells were then exposed to SJG-136 by systemic administration followed by retrieval of the hollow fibres after a suitable time period, and quantitation of the viability of the cells within using the MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazol) assay[160]. In the NCI Standard Hollow Fibre assay[158, 160], SJG-136 was evaluated against 12 different tumour cell lines, and achieved > 50% growth inhibition in 83% of the cell lines growing in fibres in the i.p. cavity, and in 29% in the s.c. location. Overall, SJG-136 produced cell kill in 5 of the 12 cell lines used in the study (*i.e.*, NCI-H522 lung adenocarcinoma, 13% kill; LOX IMVI melanoma, 10% kill; UACC-62 melanoma, 10% kill; MDA-MB-435 breast carcinoma, 29% kill; OVCAR-3 ovarian adenocarcinoma, 47% kill). During the course of the experiment, the average body weight loss in mice treated systemically with SJG-136 was ≤ 9% with no drug-related deaths, indicating that the agent was well tolerated.

Based on these results, the NCI evaluated the *in vivo* efficacy of SJG-136 in 10 human tumour xenograft models[151, 161] using tumour cell lines selected from the 60 cell line

screen and Hollow Fibre assay data[158, 160, 162]. These included LOX IMVI and UACC-62 (melanomas), OVCAR-3 and OVCAR-5 (ovarian carcinomas), MDA-MB-435 (breast carcinoma), SF-295 and C-6 (gliomas), LS-174T (colon carcinoma), HL-60 TB (promyelocytic leukaemia) and NCI-H522 (lung carcinoma). SJG-136 showed time-and dose-dependent anti-tumour activity in all of these tumour models, although the NCI-H522 and HL-60 tumours were particularly sensitive. Activity was demonstrated in all mouse models with the best efficacy observed with it being dosed once daily for 5 days, and similar results were obtained in rat xenograft models.

SF-295 and LOX IMVI were selected for assessment of maximum-tolerated dose (MTD) and minimum effective dose (MED) values in the mouse model, which were established as 120 µg/kg/day and 16 µg/kg/day, respectively. Additionally, HL-60 TB and NCI-H522 were selected for evaluation of alternative dosing schedules, such as a single dose every seven days for 3 treatments. The results showed that SJG-136 was well tolerated and highly efficacious in both models, with HL-60 TB found to be the most sensitive, and the once daily for five days schedule providing the best anti-tumour activity in both models. In addition, a single intravenous (i.v.) bolus regimen was compared with multiple daily i.v. administrations in the SF-295 and LOX IMVI tumour xenograft mouse models. Interestingly, an immediate tumour mass reduction was observed after the single dose administration (MTD for single bolus = 400 µg/kg/day), although tumour re-growth occurred within the next four days. In contrast, multiple daily administration of SJG-136 appeared to be more effective in maintaining growth suppression and/or tumour regression. Lastly, continuous infusion of SJG-136 was explored in the SF-295 and LOX IMVI xenograft mouse models but did not provide any greater efficacy compared to bolus administration.

Next, the early-stage LOX IMVI tumour xenograft model was adopted for an in-depth evaluation of the *in vivo* anti-tumour activity of SJG-136. After a single i.v. administration of 75 µg/kg, 5 out of 8 animals became tumour free by the end of the experiment (68 days), and the other three animals had a mean tumour growth delay of 26.5 days. Additional tumour growth delay experiments were carried out in the SKOV-3 ovarian model using different dose schedules including daily for five days, and every fourth day for three treatments, at SJG-136 concentrations between 25 to 120 µg/kg. In these studies, four mice experienced a partial response (PR) and three a complete

response (CR), demonstrating cumulative dose-dependent activity with all treatment schedules tolerated. SJG-136 was also evaluated in an advanced LS 174T xenograft model administered as a single-dose, daily for five days, or daily for four days for three treatments, to study dose accumulation. As for the previous models, all treatment schedules were well tolerated, and SJG-136 exhibited significant anti-tumour activity with a correlation between efficacy and cumulative dose, as well as substantial tumour growth delay and survival benefit after multiple daily administrations. A leukaemia xenograft model (HL-60) was also used to evaluate SJG-136, in which it showed dose-dependent efficacy at concentrations between 25-60 µg/kg. Therefore, SJG-136 was found to be significantly active in all four *in vivo* tumour xenograft models studied (*i.e.*, melanoma, ovarian, colon and leukaemia), and with dose-dependency after multiple daily i.v. administrations, and with greater efficacy following drug accumulation.

Finally, the efficacy of SJG-136 in the cisplatin-sensitive human ovarian cancer parental cell line CH1 was studied[158]. It was found to have significant anti-tumour activity in this model, providing tumour growth delays at a dose of 0.2 mg/kg given i.v. on days 0, 4 and 8 comparable to cisplatin at 4 mg/kg using the same schedule. More importantly, SJG-136 provided significant growth delay in the related cisplatin-resistant CH1*cis*R tumour model at 0.2 mg/kg, whereas cisplatin was inactive at 4 mg/kg[158].

In order to measure the pharmacokinetics of SJG-136 in the animal models, and to develop methodology that could potentially be used in clinical trials, a sensitive, selective and reproducible reversed phase LC/MS assay was initially developed to assess pharmacokinetic parameters in the mouse models[163]. Using this assay, it was determined that SJG-136 was detectable in plasma up to 4 hours after a single administration with a peak plasma concentration of 336 nM and a C_{max} of 30 minutes. The measured half-life ($t_{1/2}$) was 0.98 h, with a total clearance of 17.72 mL min^{-1} kg^{-1}. Overall, SJG-136 was shown to be stable in blood and plasma over a 6 hour period with no significant loss due to degradation over this time. The plasma protein binding capacity was found to be moderate at 65 ± 11% and 76 ± 5% for initial concentrations of 100 and 1,000 nM, respectively. Similar pharmacokinetic studies were carried out in a rat model following single-dose (15 and 50 µg/kg) or multiple-dose (25 µg/kg/day for 5 days) administration using a similar LC/MS/MS method[164]. In these studies, no plasma accumulation of SJG-136 was observed after administration of 25 µg/kg/day for 5 days. Another direct liquid

chromatography assay was developed to quantify the reactive imine form of SJG-136, which proved to be reproducible, linear and accurate[135]. The metabolism of SJG-136 was also evaluated using a rat liver microsome assay in which the agent was found to be metabolised by CYP3A isoforms, to have a half-life of 9 min, clearance of 190 mL/min/m^2, and a volume of distribution of 1780 mL/m^2 [164].

To study the pharmacodynamic (PD) endpoint of the mechanism of action of SJG-136 in the *in vivo* xenograft experiments, the COMET assay was used to evaluate tumour biopsies taken from the s.c.-growing LS 174T human colon cancer xenograft model in MF1 nude mice[165]. SJG-136 caused a significant growth delay in this tumour model when delivered i.v. at 0.30 or 0.45 mg/kg, and tumour samples were removed at 1, 3 and 24 hours, and then subjected to COMET analysis. A small but significant level of DNA cross-linking was observed at 1 hour which remained constant over 24 hours. However, the extent of cross-linking increased with the higher dose of 0.45 mg/kg which also remained constant over 24 hours[158].

Finally, an *ex vivo* bone marrow (CFU-GM) colony formation assay was used to assess the toxicity of SJG-136 toward the hematopoietic system[158, 166]. The results demonstrated that the concentrations of SJG-136 required to achieve cytostatic and/or cytocidal effects in HL-60 TB and Molt-4 human leukaemia tumour cells in cell culture or in a soft agar colony formation assay (*i.e.*, the IC_{50}, IC_{75}, and IC_{90} values) were much lower than the concentrations required for similar effects in an *ex vivo* bone marrow assay[158], thus suggesting a significant selective cytotoxicity toward the tumour cells.

In summary, SJG-136 showed significant activity in most of the *in vivo* models examined after short exposure times at nanomolar concentrations, or for longer exposure times at sub-nanomolar concentrations. For example, in the early-stage NCI xenograft studies, significant tumour growth delays were observed in seven of the eight models after i.v. administration, with multilog cell kill in four of the eight models. It was also found to be markedly efficacious in two advanced-stage xenograft models (LOX IMVI melanoma and HL-60 TB promyelocytic leukaemia) with similar results to those obtained in the early-stage models. Overall, it was established that SJG-136 is significantly active after a single bolus administration but provides the best efficacy after administration multiple times over a five-day period[151].

1.6.3 Clinical evaluation of SJG-136

The first clinical evaluation of SJG-136 as a single agent in a Phase I setting was carried out in patients with advanced and/or metastatic solid tumours refractory to conventional anticancer therapies, or for whom no conventional treatment was available[98b]. The purpose of the study was to establish the Maximum Tolerated Dose (MTD) along with pharmacokinetic and pharmacodynamic profiles using COMET[153] and γ-H2AX[155] foci assays. SJG-136 was administered i.v. every 21 days at different doses. An MTD of 45 μg/m^2 was established for this 21-day schedule. The major observed drug-related adverse events were reversible liver toxicity (*i.e.*, transaminitis), fatigue and a delayed lower-limb oedema. Out of twelve evaluable patients, stable disease (SD) was observed in ten patients, but no partial (PRs) or complete responses (CRs) were reported.

A second Phase I clinical trial with a different dosing schedule was carried out in patients with refractory solid tumours using a starting dose of 10 μg/m^2 per day. In this case, SJG-136 was administered as a bolus infusion over 20 minutes on days 1, 8 and 15 of a 28-day cycle. The MTD was assigned as 40 μg/m^2 per day.

The third Phase I clinical trial was carried based on the previous Phase I clinical study conducted by Hochhauser and co-workers[98b]. This study evaluated two different dose schedules. Sixteen patients were enroled, and a MTD of 30 μg/m² was determined. The DLTs observed were similar to those documented in the previous Phase I clinical trials (*i.e.*, oedema, fatigue, dyspnoea and elevations in liver transaminases and alkaline phosphatase). In this trial, one patient experienced a confirmed partial response (PR), and one an unconfirmed PR. Furthermore, two patients achieved stable disease (SD) lasting more than 12 weeks.

On the basis of these encouraging Phase I results, SJG-136 progressed to a multicentre Phase II clinical evaluation in patients with epithelial ovarian cancer who had not responded to prior treatment with cisplatin or carboplatin. Nineteen patients were enroled with SJG-136 administered i.v. over 20 minutes on days 1-3 every 21 days. The goals of the study were to evaluate the overall response rate to SJG-136, and to assess toxicity. Furthermore, for each patient the extent of adduct formation was determined in peripheral blood mononuclear cells (PBMCs) and tumour biopsies using the COMET and γ-H2AX

assays. However, patient recruitment to this trial was poor, and no clinical results have been reported to date.

A second Phase II clinical evaluation of SJG-136 in patients with advanced chronic lymphocytic leukaemia (CLL) and acute myeloid leukaemia (AML) has been terminated in November 2015. In this trial patients received SJG-136 intravenously at a dose level of 15 μg/m^2/day (*Cohort I*) or 30 μg/m^2/day (*Cohort II*) on days 1, 2 and 3 every 21 days for six cycles. The maximum tolerated dose (MTD) was assessed based on the DLTs observed and, for secondary outcome measures the safety profile of SJG-136 was evaluated based on pharmacokinetic parameters such as maximum plasma concentration (C_{max}), time to reach C_{max} and terminal half-life ($t_{1/2}$) accompanied by haematology and serum chemistry.

1.7 Methods used in this study to study the interaction of small-molecule binding ligands with oligonucleotides

This section will focus on the principles and background of the methods and techniques used in this study. The biophysical study applied reversed phase high performance liquid chromatography (RP-HPLC), matrix-assisted laser desorption/ionisation time of flight mass spectrometry (MALDI-TOF-MS), circular dichroism (CD), fluorescence resonance energy transfer (FRET) and molecular modeling (MM) studies. Moreover, the traditional polymerase chain reaction (PCR), quantitative polymerase chain reaction (qPCR) and Western blotting (WB) were used during the biological analysis of SJG-136.

1.7.1 Reversed Phase High Performance Liquid Chromatography (RP-HPLC)

RP-HPLC is a very powerful and widely-used technique for the separation, purification and analysis of complex mixtures comprising large and small biomolecules, such as proteins, peptides and nucleic acids[167]. In particular, it is a useful analytic technique to study the interaction of small-molecules with specific DNA target sequences of oncogenic transcription factors. Preliminary data can easily and rapidly be obtained on the ability of ligands to bind to consensus transcription factor sequences. In the case of ligand/DNA interaction, additional peaks appear in the chromatographic profile due to a change in the polarity of the ligand/DNA complex compared to the DNA alone.

The principle of RP-HPLC is an adsorption mechanism which depends on hydrophobic interactions between solubilised molecules (solutes) in a mobile phase and specifically modified ligands chemically fixed to a non-polar chromatography matrix (sorbent, stationary phase). The sample to be analysed is introduced into the mobile phase solution with a high pressure pump by injection and subsequently forced through the column (HPLC tube, stationary phase) (**Figure 1.29**). The solute separation occurs within the column based on hydrophobicity. Non-polar molecules contained in the mixture form strong attractions with the stationary phase through van der Waals interactions, and therefore will distribute on the surface of the stationary matrix. In contrast to this, polar components lack the hydrophobic association with the stationary phase and consequently, will remain in the mobile phase and migrate through the column more rapidly. The

absorbance of eluted components is monitored using a UV detector which is placed directly after the column. The intensity of the UV signal depends on the amount of a particular compound in the mixture passing the HPLC tube and absorbing light. The UV signals are converted into electrical signals appearing as peaks in the chromatographic profiles. By measuring the area under the peak (AUC) precise quantification is possible as the AUC is proportional to the concentration of the analytes.

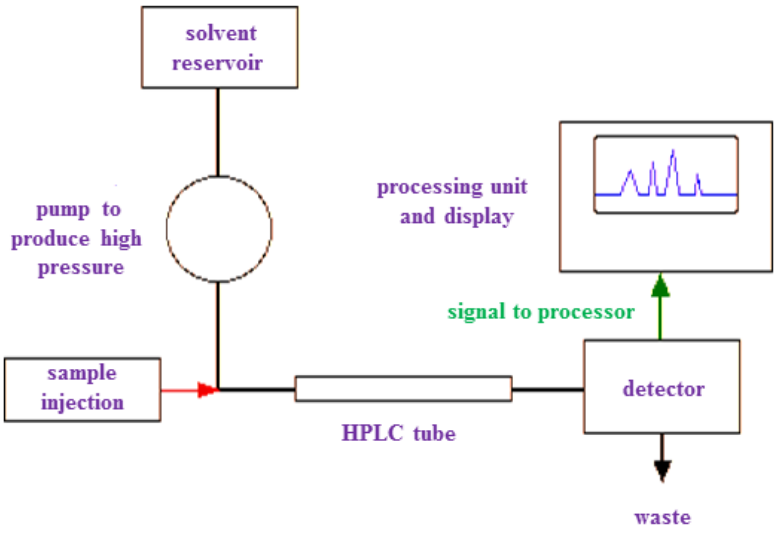


Figure 1.29: Reversed phase HPLC separation: Overview. Adapted from *www.chemguide.co.uk/analysis/chromatography/hplc.html*.

The stationary phase of RP-HPLC consists of hydrophobic ligands which are incorporated into a porous, insoluble beaded matrix. Silica is the most commonly used support material in RP-HPLC[168]. The surface active sites on the silica gel, the silanol groups (Si-OH) are modified by attaching long non-polar linear hydrocarbon chains (C1, C4, C6, C8, C18, *etc*.) or phenyl and cyanopropyl ligands in order to increase the hydrophobicity of the stationary phase (**Figure 1.30**). The matrix plays a fundamental role during the separation process as the interaction of the solutes with the stationary phase occurs on its surface and is highly dependent on the type of bonded phase, particle size, pore size and alkylchain density as well as its length[169]. The precise mechanism of retention is still an on-going debate[170]. It is presumed to be governed by a combination of adsorption and a partitioning type mechanism where the solute molecules embed fully between the hydrocarbon chains[171].

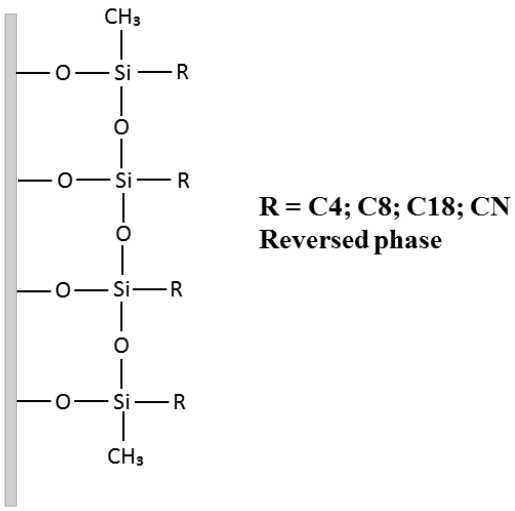


Figure 1.30: The surface of reversed phase stationary phases.

The mobile phases used in RP-HPLC are mixtures of water and organic solvents (modifiers), typically acetonitrile or methanol. The role of the organic modifiers is to compete against the ligands of the stationary phase for the attraction of the solute molecules. Initially, the analysed mixture is introduced into the system with the mobile phase containing low concentrations of organic modifiers with the aim to allow non-polar solutes to adsorb to the hydrophobic matrix. During the RP-HPLC process, the concentration of the organic modifier is gradually increased to favour desorption of the solutes from the reversed phase medium back into the mobile phase. Consequently, very polar components require low concentrations of organic modifiers, migrate through the stationary phase more quickly, and hence elute from the column prior to less-polar molecules. Hydrophobic solutes will form strong attractions with the matrix ligands through van der Waals interactions, and thus spend more time in the column.

The solutes separation can proceed either by an isocratic (concentration of organic modifier is constant during the HPLC process) or gradient elution where the concentration of the organic solvent is gradually increased from high to low polarity.
Ion pair RP-HPLC is the predominant chromatography mode applied for the study of nucleic acids, including DNA and RNA[172]. The aqueous mobile phase often contains ionic additives, referred as ion pair agents. The ion pairing reagents are cationic amines, typically triethylammonium acetate (TEAA) or triethylammonium bicarbonate (TEAB), which form ion pairs with the negatively charged phosphate groups of the nucleic acids

making them electrically neutral. This process increases the hydrophobic character of the nucleic acids and allows the separation of ionic and highly polar molecules on a reversed phase column, such as DNA, RNA as well as hydrophilic peptides.

1.7.2 Matrix-Assisted Laser Desorption/Ionisation Time of Flight Mass Spectrometry (MALDI-TOF-MS)

MALDI is the most popular analytical technique for analysing low and high molecular weight compounds, particularly oligonucleotides, peptides, proteins and oligosaccharides[(198-200)]. MALDI-TOF is widely used in the analysis of drug-DNA interactions. It precisely determines the mass of formed ligand-DNA complexes[173]. This analytical application allows rapid and accurate mass identification with detection sensitivity in the picomolar concentration range[174].

A typical MALDI instrument consists of three essential components (**Figure 1.30**): ion source, mass analyser and detector. The analysed mixture is embedded in a particular matrix which is exposed to a focused laser beam, commonly nitrogen laser pulse at 337 nm. The crystallised matrix components absorb the laser energy, followed by an ionisation process which in turn results in charge transfer from the matrix molecules to the analysed compounds. Subsequently, the charged molecules pass an electric field of known strength and are separated in the mass analyser according to their molecular weight (time of flight). Molecules with higher molecular weight fly slowly through the mass tube compared to compounds with low molecular mass, thus reaching the detector at a later time point. The flying speeds of the produced ions are proportional to their mass-to-charge ratio. The detector counts the numbers of different ions and generates the mass spectrum which can be recorded in positive as well as negative mode.

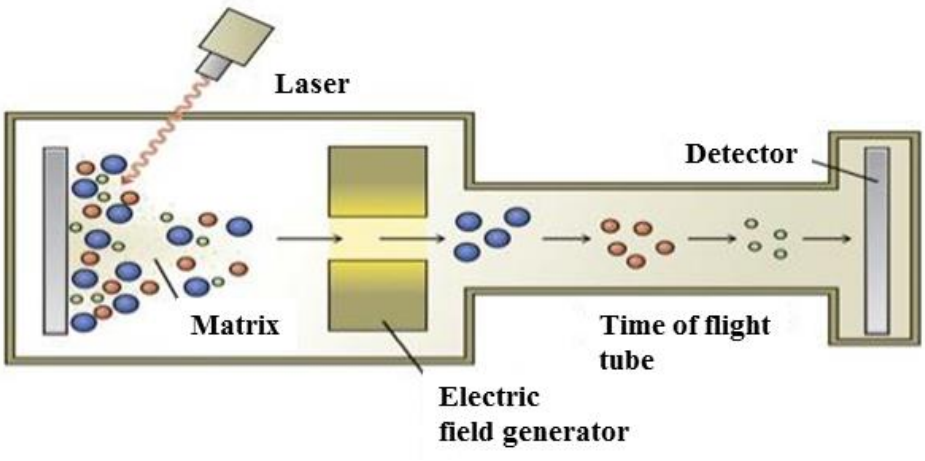


Figure 1.30: Principle of MALDI-TOF. Adapted from *www.studyblue.com*.

Two different types of analyser are available for MALDI-TOF-MS: the linear mode and the reflectron mode analyser. The linear mode analyser simply measures the required time for an ionised compound to fly from one end of the mass analyser to the detector. During the reflectron mode, the charged molecules initially pass a linear mass analyser and are then reflected, followed by passage through a second linear mass analyser before they reach the detector. The main advantage of the reflectron mode is the fact that the charged molecules spend more time flying, allowing an improved separation of molecules with very similar mass-to-charge ratios yielding a higher resolution.

The choice of matrix plays a fundamental role in MALDI-TOF-MS, affecting sensitivity and reproducibility of mass spectrometry analysis. Benzoic acid, cinnamic acid and related aromatic compounds are commonly used for the analysis of proteins[175]. 3-Hydroxypicolinic acid[176] and succinic acid[177] demonstrated to be suitable matrices for oligonucleotides. The appropriate matrix is combined with the analyte of interest by solubilisation in a mixture of water and an organic solvent in order to ensure that hydrophobic as well as hydrophilic components completely dissolve. The matrix mix is then spotted onto a MALDI target plate and dried naturally for crystallisation. Finally, the laser pulse is applied resulting in matrix ionisation and charge transfer to the analysed components (**Figure 1.31**).

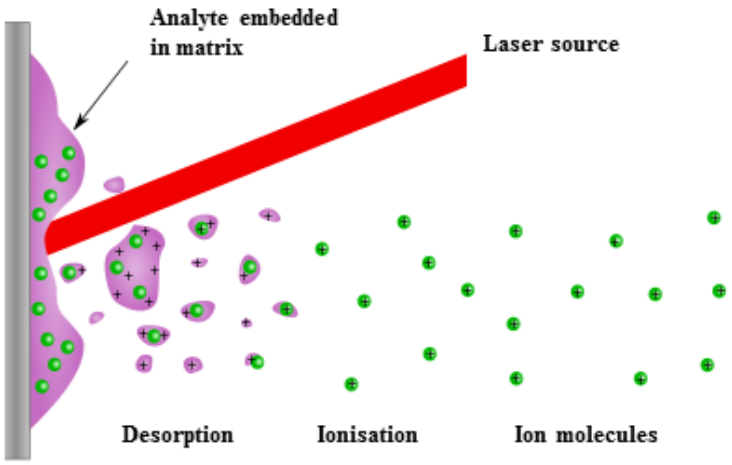


Figure 1.31: Principle of ionisation process during MALDI. Adapted from *fr.wikipedia.org*.

1.7.3 Circular Dichroism (CD)

CD is a powerful spectroscopic method used to study conformational changes in the secondary structure of oligonucleotides[178]. It is an indispensable technique for the investigation of the various conformational forms of nucleic acids, such as A-form, B-form, Z-form, quadruplexes and other structures[178], hence a valuable tool for mapping conformational features of specific oligonucleotide sequences. More importantly, CD can be applied to evaluate conformational changes in the secondary structure of DNA upon binding of specific DNA-interacting ligands. Therefore, in context to this project, CD is a useful tool that may provide important data on anti-tumour agents that specifically target and regulate fundamental elements crucial for gene expression which is also the focus of current anti-cancer drug development[178c].

CD originates from the interaction of chiral molecules (vast majority of biological molecules) with circularly polarised light which is the sum of left-handed and right-handed linearly polarised light[179]. Chiral molecules absorb left-handed and right-handed light differently and the resulting difference is simply the CD signal, called ellipticity θ and expressed in degrees. Ligands bound to biological compounds alter its secondary structure leading to a change in characteristic CD signals of the native molecules providing information about binding constants, kinetics of binding as well as conformational changes in the ligand/molecule complex.

A monochromator produces linearly polarised light which subsequently converted to circularly polarised light by a photoelastic modulator (PEM) (**Figure 1.32**). Following this, the circularly polarised light beam passes through the optically active sample where it is absorbed differently. The photomultiplier detects the difference in the absorbed transmitted light and measures the exhibited CD signal. CD studies of oligonucleotides mainly employ UV within 180-300 nm where the DNA bases absorb light. $\pi \rightarrow \pi^*$ transitions in purines and pyrimidines have been shown to be responsible for the characteristic CD bands of nucleic acids[180]. Of utmost importance is that the position and amplitude of CD signals of various DNA conformations, such as A-form, B-form, Z-form, hairpin or quadruplex strongly depend on the base pair sequence[181].

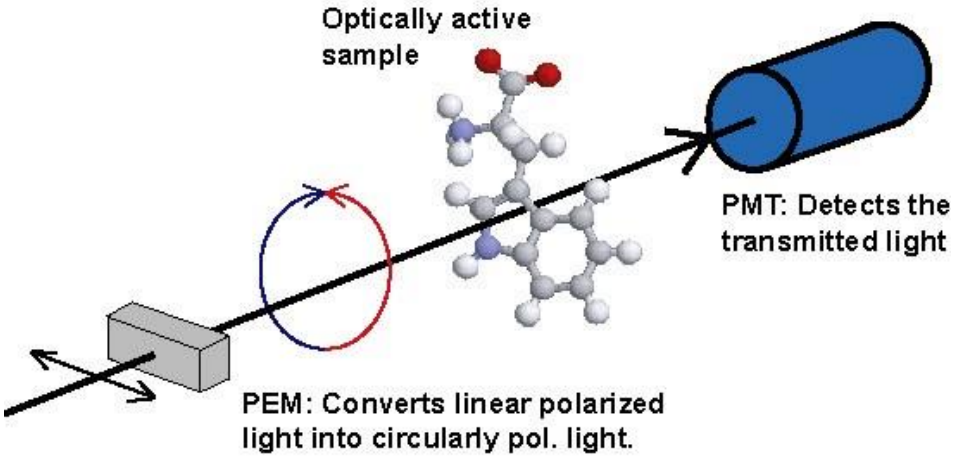


Figure 1.32: The principle behind CD spectroscopy. The light from UV is passed into a Photo Elastic Modulator (PEM) which converts the linearly polarized light into alternating left and right handed polarized light. The two polarizations are differently absorbed, and the difference in absorption is detected with a Photo Multiplier Tube (PMT). Taken from *www.isa.au.dk*.

CD spectroscopy is a fast and relatively inexpensive technique. It exhibits high sensitivity at low DNA concentrations (20 µg/mL). Furthermore, short oligonucleotides as well as longer DNA fragments up to 8,000 base pairs can be easily analysed using CD[182]. And lastly, CD is capable of examining solutions and films[183]. However, CD cannot provide any structural information of the analysed molecules on the atomic level in contrast to X-

ray diffraction and NMR spectroscopy, hence CD is primarily utilised for preliminary studies.

1.7.4 Fluorescence Resonance Energy Transfer (FRET) assay

FRET is a distance-dependent physical phenomenon in which energy from an excited donor molecule is transferred to a suitable acceptor molecule in the absence of radiation[184]. FRET is widely applied to investigate molecular interactions in biomedical research and drug discovery. FRET data can provide useful information about structure and conformation of proteins[185] and nucleic acids[186]. Small-molecules stabilise the secondary structure of DNA upon binding which results in an increase in the melting temperature of the ligand/DNA complex. The differences in the melting temperature of the DNA/ligand complex are compared with the melting temperature of the DNA alone and expressed as ΔT_m and can provide useful data about the reactivity of ligands toward specific DNA sequences.

The donor probe, which is always a fluorescent molecule, absorbs energy resulting in excitation from its ground state to an excited state within picoseconds[187]. If an acceptor molecule is located in close proximity, the emitted energy from the excited donor molecule is subsequently transferred to the acceptor molecule through long-range dipole-dipole interactions (**Figure 1.33**)[184]. This process leads to a decrease of the fluorescence intensity of the donor molecule and an increase in the emitted fluorescence of the acceptor molecule which is monitored by a UV/VIS detector.

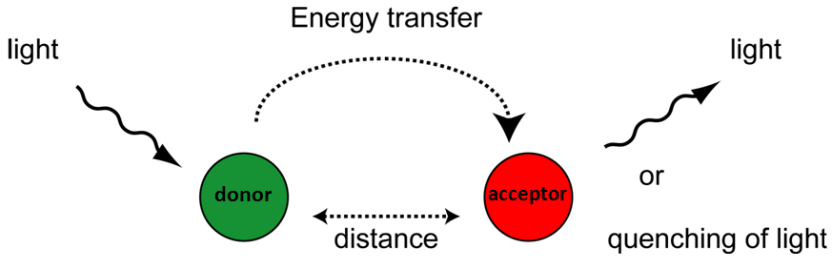


Figure 1.33: Principle of the FRET assay. Taken from *www.molecular-beacons.org*.

The process of energy transfer occurring in our study is an inversion of the FRET principle described above. The hairpin oligonucleotides were tagged on the 5'-end with the fluorescent donor molecule 6-carboxyfluorescein (FAM) and on the 3'-end with the acceptor probe 6-carboxytetramethylrhodamine (TAMRA) (**Figure 1.34**). TAMRA owns a quencher role and neutralises the emitted fluorescence of the donor probe due to the close proximity of both molecules. An increase in the melting temperature results in destabilisation of the hairpin leading to a conversion to linear single-strand. Consequently, this reaction removes the proximity of donor and acceptor and allows the fluorescent probe to emit light (**Figure 1.34**).

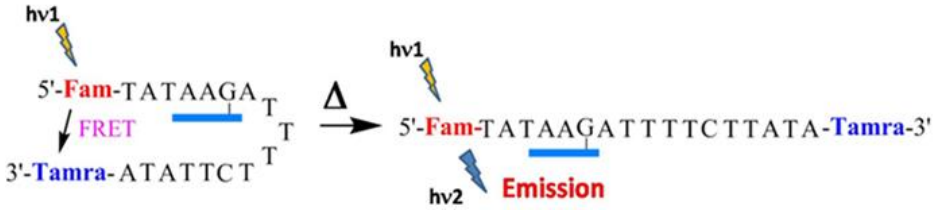


Figure 1.34: The principle of the FRET assay used in this study.

1.7.5 Molecular modeling

Molecular modelling techniques allow for the atomic-level evaluation of a molecule's shape-fit and binding affinity with a receptor. These techniques find use in establishing structure-activity relationships of libraries of ligands, thereby playing a role in the drug discovery process. The vast majority of molecular modelling techniques involve the use of static structures derived through crystal or NMR studies. Molecular dynamics (MD) simulations, on the other hand, consist of the calculation of the time-dependent behaviour of molecular systems, and, therefore, involve simulation of protein or DNA. MD simulations are far more robust and reliable than modelling analysis alone, and have provided valuable information on the changes in conformations of biomolecules (for example proteins or nucleic acids) as predicted over a certain time-course.

The field of molecular dynamics simulations has progressed rapidly through technological advancement. The speed of computers has doubled every two years ('Moore's Law'[188]) for the past few decades, and the use of specialised GPUs (Graphics

Processing Unit) and super-computers has allowed the simulation of large macromolecules. This has culminated in the 1 ms all-atom molecular dynamics simulations of BPTI (Bovine Pancreatic Trypsin Inhibitor)[189] using a specially built, massively parallel super computer capable of simulating in orders of magnitude greater than standard computer systems[190]. Statistical mechanics is central to the concept of molecular dynamics simulations and it is the study of microscopic systems to produce macroscopic conclusions. In MD simulations, individual atoms within systems are simulated in order to determine features of the system. Ensembles are defined as systems which have different microscopic states, but identical macroscopic ones, and in order to simulate whole systems, averages corresponding to experimental observables are defined in terms of 'ensemble averages'. The central hypothesis to this is that over the course of a simulation, all possible macroscopic states should be simulated by different microscopic states, and, therefore, at the end of a simulation, the system will have passed through all states. From a drug discovery perspective, the simulation of a ligand with a receptor should result in the determination of the lowest energy conformation (and thus, most biologically likely) combined drug:receptor complex structure.

In the case of this study, a molecular dynamics-driven approach was undertaken, as an MD protocol has been developed and extensively validated in this lab for use with pyrrolobenzodiazepine-based DNA-drug adducts[139, 191]. The protocol involves the simulation of each PBD (*i.e.*, SJG-136, GWL-78 and anthramycin) both covalently and non-covalently bound to DNA over a 10ns time-course in implicit solvent. Non-covalently bound simulations are used to visually examine the shape-fit of the molecule in DNA and to calculate affinity of the molecule for a particular sequence. This is accomplished through free energy of binding calculations. Covalently bound simulations are then used to both investigate degree of DNA disorder once the PBD is bound to DNA, and also to ascertain the most preferred guanine for alkylation through potential energy calculations. Using this combined approach, it has been possible to fully rationalise results of biophysical assays undertaken in this study.

1.7.6 Polymerase Chain Reaction (PCR)

PCR is a quick, inexpensive and simple tool that allows the production and detection of a specific DNA segment more than billion-fold from a complex pool of DNA. This

scientific technique is routinely used in the clinic for diagnostic purposes and in the field of molecular biology for a variety of applications[192]. Popular applications for PCR are the detection of hereditary diseases, identification of genetic fingerprints in forensic science, diagnosis of infectious diseases (*e.g.*, HIV, hepatitis, mycoplasma or pneumonia), DNA cloning Southern blotting, DNA sequencing and recombinant DNA technology. PCR is also a valuable tool for gene sequencing, particularly those implicated in the development and progression of human neoplasms and this has relevance to this body of work. In context to the body of this work PCR was used to analyse the changes in gene expression profiles before and after treatment with the ligand.

Four primary components are essential for PCR: a DNA sample, gene-specific primers, DNA building blocks and the DNA polymerase. The DNA sample serves as the template and contains the DNA sequence of interest to be amplified, obtained from a variety of tissues and organisms, such as in *vitro* cultivated cells or peripheral blood. Even trace amounts of DNA samples are sufficient to generate enough copies. Primers are short (typically 18-25 base pairs), sequence-specific oligonucleotides generated *via* chemical synthesis. They are sequence-specific and complementary to the DNA region of interest. The DNA building blocks are composed of the nucleotides adenine, guanine, thymine and cytosine that are crucial for the DNA polymerase in order to create the new DNA strand. Lastly, the DNA polymerase is the key element of the PCR. The thermostable Taq polymerase can withstand high temperatures required for the PCR process[193]. It assembles the single bases to a concise DNA strand on the basis of the DNA template.

The PCR process consists of three consecutive steps (**Figure 1.35**). In the first step, also referred to as denaturing, the DNA sample is heated to temperatures between 90 and 97°C leading to breakage of the hydrogen bonds between the double-stranded DNA and separation into two single strands (**Figure 1.35**). The second step, the annealing/hybridisation, involves sample cooling to usually 50 °C which enables the specific primers to associate with the complementary DNA sequences (**Figure 1.35**). During the last step (the extension) the Taq polymerase adds the single nucleotides to the ends of the primers at 72 °C, thus generating a sequence specific complementary DNA strand (**Figure 1.35**). Repetition of these three steps results in doubled copy number with

each cycle. However, increasing the cycle number above 35 has a little positive effect due to depletion of the reagents.

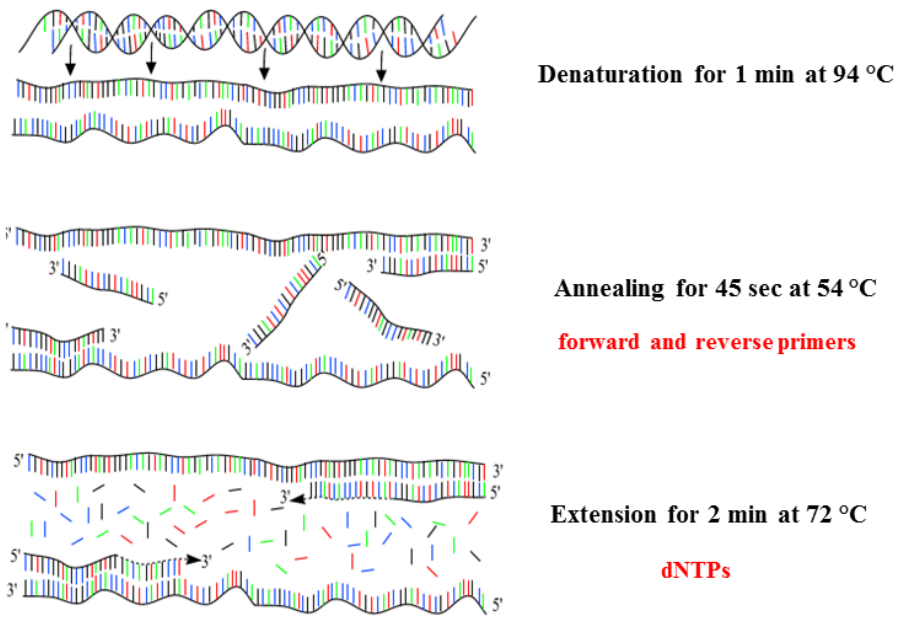


Figure 1.35: Principle of the polymerase chain reaction. Adapted from *http://users.ugent.be/~avierstr/principles/pcr.html*.

The traditional method applied for PCR product visualization and analysis is agarose gel electrophoresis. The products are stained with DNA-intercalating chemicals, such as ethidium bromide followed by separation on an agarose gel. Smaller DNA fragments migrate through the gel toward the anode faster than larger strands. A DNA ladder which represents a predetermined set of DNA products with known sizes is injected onto the agarose gel simultaneously allowing size comparison with the PCR products.

1.7.7 Quantitative Polymerase Chain Reaction (qPCR)

Despite the significant benefits and advantages, the traditional PCR, often termed as end-point PCR, has some limitations. The term 'End-point' simply means that the detection and quantification of the amplified DNA sequence are undertaken once the reaction is complete after the last PCR cycle using agarose gel electrophoresis and image analysis. The DNA-intercalating dye ethidium bromide lacks sensitivity and consequently binds to any double-stranded DNA present in the sample[194]. Moreover, the resulting end-point

levels of PCR products vary strongly from PCR run to run leading to inconsistent results[195]. Additionally, image analysis is dependent on the brightness of the produced bands which also differ from experiment to experiment. These drawbacks led to a refinement of the traditional PCR by Higuchi and co-workers in 1992[192]. The development of quantitative PCR (often referred to as real-time PCR) allows simultaneous amplification and detection of specific DNA sequences while the amplification is occurring, in 'real-time', utilising fluorescent technology (fluorescent reporters). The main advantage of qPCR is its ability to precisely and accurately determine the initial amount of target DNA contained in a complex mixture even if only a few DNA copies are present. The most common applications for qPCR are pathogen detection, gene expression profiling, single nucleotide polymorphism analysis, analysis of chromosome aberrations and also recently protein detection by real-time immune PCR[196].

The basic principle of qPCR is the introduction of fluorescent reports into the sample that bind to the target DNA resulting in the generation of a fluorescence signal. The fluorescence intensity is directly proportional to the amount of double-stranded and increases during the amplification progress as the DNA sequence of interest accumulates. Hence, the changes in fluorescence signal are monitored over the course of the reaction by an instrument that combines thermal cycling with fluorescence measurement.

SYBR green I, an asymmetric cyanine dye, emits 1,000-fold greater fluorescence when it binds to the minor groove of DNA compared to its free from in solution[197]. Similar to ethidium bromide, SYBR green I lacks selectivity as it interacts with any present double-stranded DNA.

An alternative approach is the use of hydrolysis probes that are sequence-specific dually fluorophore-labelled short oligonucleotides[198]. Hydrolysis probes contain a fluorescent reporter molecule at the 5' end and a quencher molecule at the 3' end of the molecule, thereby forming a donor-acceptor pair[199]. In case of close proximity of the reporter and the quencher molecule, the quencher absorbs the fluorescence emission of the reporter based on the FRET principle as described in Section 1.7.4. Following attachment of the hydrolysis probe to the target DNA, it is cleaved from the DNA strand by the Taq polymerase due to its 5'-exonuclease activity during the amplification (**Figure 1.36**). As

a result, the reporter and the quencher get separated allowing the reporter's energy and fluorescence to be liberated (**Figure 1.36**). Hydrolysis probes offer greater specificity because they bind specifically to the DNA segment of interest, thus only sequence-specific amplification is captured during the PCR reaction.

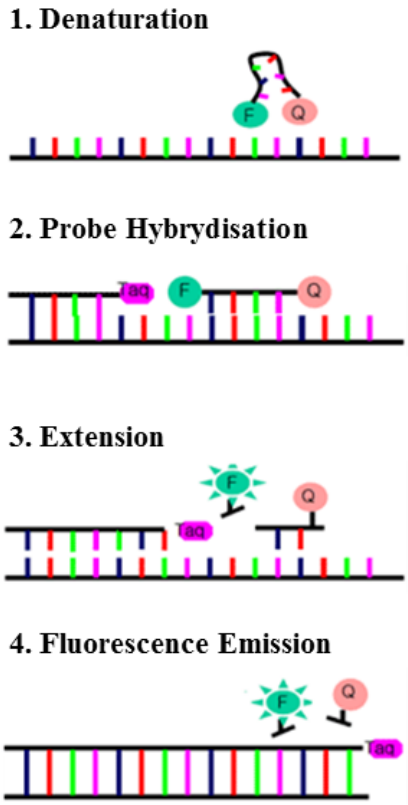


Figure 1.36: Hydrolysis probe detection during qPCR. Adapted from *http://eng.bioneer.com/products/geneexpression/qPCRArrayService-detection.aspx*.

1.7.8 Western blotting

Western blotting, also known as immunoblotting, is a well-established and widely accepted analytical technique used in molecular biology, biochemistry and immunogenetics to separate and identify specific proteins from a complex protein mixture. Proteins can be detected in ng-range due to the high resolution of the electrophoresis and the high specificity and sensitivity of the immunoassay. This technique is of utmost importance for the investigation of ligands that specifically inhibit gene expression. As protein synthesis is the final process from transcription to translation, examination of protein profiles before and after treatment with ligands enables to

determine whether the DNA-binding agents indeed bind to the targeted sequences alter the gene expression into a favourable direction.

The Western blotting procedure is composed of four major steps: SDS-PAGE, transfer onto the membrane, probing with antibody and detection. The complex protein samples originate typically from cell or tissue homogenates that first have to undergo preliminary treatment before continuing to separation by gel electrophoresis. Reducing agents, such as 6-mercaptoethanol are applied to the protein samples in order to disrupt the quaternary and tertiary protein structures by breaking any inter-and intra-chain disulfide and to linearise the proteins which is a crucial step for subsequent SDS-PAGE. Moreover, SDS, an anionic detergent is added to the protein samples with the aim to mask any inherent charge by negatively charged detergent micelles, also an essential process for SDS-PAGE. Followed this, the protein samples are subjected to SDS-PAGE where they migrate toward the anode due to their negative charge and separate according to physical properties, including size, charge and structure. After electrophoresis the proteins are transferred onto a membrane, usually nitrocellulose or polyvinylidene difluoride (PVDF). A gel-membrane-filter sandwich is assembled with the membrane facing the gel surface and the anode so that the negatively charged proteins located on the gel travel through the membrane toward the anode when an electric current is applied. Subsequently, the membrane is probed with a primary antibody that specifically recognises and binds to proteins or epitopes on proteins. The unbound primary antibody is washed off the membrane and an enzyme labelled secondary antibody (*e.g.*, horseradish peroxidase, HRP) is incubated with the membrane that in turn precisely interacts with the bound primary antibody (**Figure 1.37**). Addition of substrate molecules results in a reaction between the enzyme bound to the secondary antibody and the substrate generating coloured bands (**Figure 1.37**). Detected signals represent proteins bound to the membrane and can be further determined by their location on the blot and densitometry.

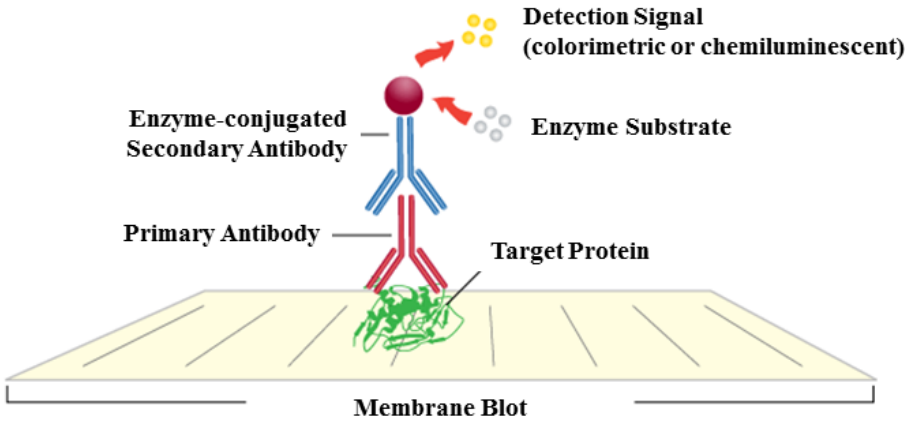


Figure 1.37: Graphical representation of detection during Western blotting. Adapted from *http://employees.csbsju.edu/hjakubowski/classes/ch331/Techniques/TechElectrophoresis.htm*.

1.8 The aims of the project

The main objective of this study was to investigate whether PBD dimers like SJG-136 also interact with transcription factor binding sequences and exert their anti-tumour activity through transcription factor inhibition in addition to previously reported mechanisms. The aim was to develop an ion pair reversed phase HPLC/MS analytical methodology in order to obtain both kinetic and thermodynamic data about the adduct formation between the PBD dimer SJG-136 and the consensus sequences of the oncogenic transcription factors NF-κB, EGR-1, AP-1 and STAT3. In addition, circular dichroism, fluorescence resonance energy transfer DNA binding assay and molecular modeling simulations were used to support the HPLC/MS observations. The biological study utilised the traditional polymerase chain reaction, quantitative polymerase chain reaction and Western blotting methods.

Chapter 2: Development of an ion pair reversed phase HPLC/MS based methodology to study the interaction between PBD monomers and dimers with consensus sequences of oncogenic transcription factors

2.1 Background

The pyrrolobenzodiazepine (PBD) dimer SJG-136 has been recently used for the development of a combined HPLC/MS assay to study the interaction of DNA minor-groove binding agents with short oligonucleotides[157]. Traditional methods applied to date for evaluating the interaction between anti-cancer agents and naked DNA are mainly gel electrophoresis-based[148]. In this method, radiolabeled, drug-treated duplex DNA is denatured by heat, and electrophoresed on a neutral agarose gel. This process allows re-annealing of the cross-link to duplex DNA induced by the cross-linking agent. The resulting adducts are then quantitated by densitometry. However, this robust and reproducible technique has many disadvantages, *e.g.* it is labour-intensive as radiolabelling, gel electrophoresis, imaging and densitometry steps are required. Also, no data relating to the rate and extent of adduct formation can be obtained from using this technique. In addition, a further drawback is the use of expensive and potentially harmful radioactivity which is necessary to label the DNA. Therefore, the main aim of this study was to develop a novel ion pair reversed phase HPLC (RP-HPLC) and MS based method which would allow the characterisation and quantitation of adducts formed between cognate sequences of oncogenic transcription factors and covalent DNA minor groove binding agents such as the PBD monomers and dimers. In contrast to the traditional method, this analytical technique would require only small quantities of short, inexpensive oligonucleotides without any use of radioactivity. Furthermore, formed adducts would be rapidly visualised (< 30 min) by HPLC and positively identified by MS. A major advantage of this newly developed technique is that both thermodynamic and kinetic data can be obtained relating to the rate of reaction for any given DNA sequence. Materials and methods used in the previously reported method are described elsewhere in detail[157].

2.2 Objectives

Based on the above reported HPLC/MS analytical method (see Section 2.1), the aim of the study was to further modify this methodology so that it can be applied to study the interaction of PBD dimers such as SJG-136 with cognate sequences of oncogenic transcription factors (TFs) (**Figure 2.1**). For NF-κB, two possible consensus sequences were investigated (NF-κB-1 and NF-κB-2).

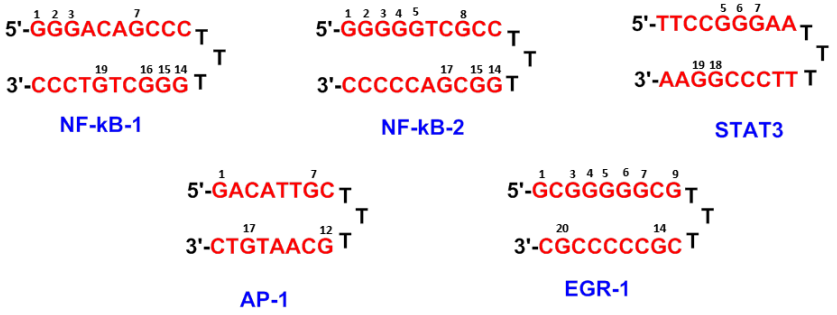


Figure 2.1: Hairpin oligonucleotides used in this study with the cognate sequences of the TFs NF-κB-1, NF-κB-2, EGR-1, AP-1 and STAT3 highlighted in red.

2.3 Methodology approach

The flow chart shown in **Figure 2.2** summarises the overall approach. Initially, it was aimed to develop an ion pair reversed phase HPLC method and validate it by investigating the interaction of the PBD monomers GWL-78 (**2**, **Figure 1.20**) and KMR-28-39 (**3**, **Figure 1.20**) with the cognate sequences of the oncogenic TFs NF-κB, EGR-1, AP-1 and STAT3 (**Figure 2.1**). Similarly, a MALDI-TOF-MS method was aimed to be developed in order to identify ligand/DNA adducts. Upon successful validation, the RP-HPLC method will be used to study the interaction of SJG-136 with the oncogenic transcription factor cognate sequences, NF-κB, EGR-1, AP-1 and STAT3, followed by identification of the formed adducts by MALDI-TOF-MS.

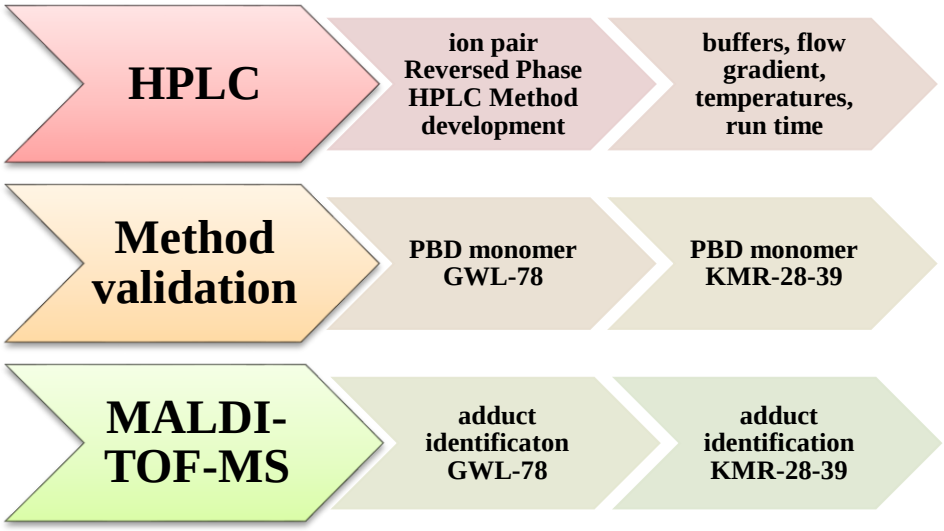


Figure 2.2: Flow chart of methodology approach.

2.4 Method development

In order to examine the reaction between PBD monomers GWL-78 (**2**, **Figure 1.20**) and KMR-28-39 (**3**, **Figure 1.20**) with NF-κB, EGR-1, AP-1 and STAT3 consensus sequences (**Figure 2.1**), an ion pair RP-HPLC method was developed as described below.

2.4.1 RP-HPLC method development

Chromatography was performed on a Dionex UltiMate 3000 system from Thermo Scientific. The instrument was equipped with a 2.1 x 50 mm XBridge™ OST C_{18} column packed with 2.5 µm particles. The general settings used are listed in **Table 2.1**. The mobile phase consisted of 100 mM triethylammonium bicarbonate (TEAB) as buffer A and 40% acetonitrile in water as buffer B. The pump gradients used for HPLC analysis are described in the individual sections.

Table 2.1: Generals setting used for the RP-HPLC study.

Parameter	Setting
Run time	25 min
Column	20 °C
Sampler	15 °C
UV	254 nm
Flow rate	0.5 mL/min

Method I

Initially, Method I was designed using the mobile phase composition listed in **Table 2.2** and **Figure 2.3,** and the instrument settings described in **Table 2.3**. The preparation procedure of hairpin DNA working solutions is described in Section 7.2.1.2. Following this, the DNA samples were subjected to ion pair RP-HPLC as described below. Based on previous studies with PBDs and 10-30 mer oligonucleotides[157] the DNA was expected to elute between 7 and 18 min corresponding to 30-45% mobile phase B.

Table 2.2: Mobile phase composition used for Method I. Phase A: 100 mM TEAB; Phase B: 40 % acetonitrile.

Time [min]	A [%]	B [%]
0	90	10
5	80	20
7	70	30
18	55	45
20	30	70
22	20	80
23.5	10	90
24	10	10
25	90	10

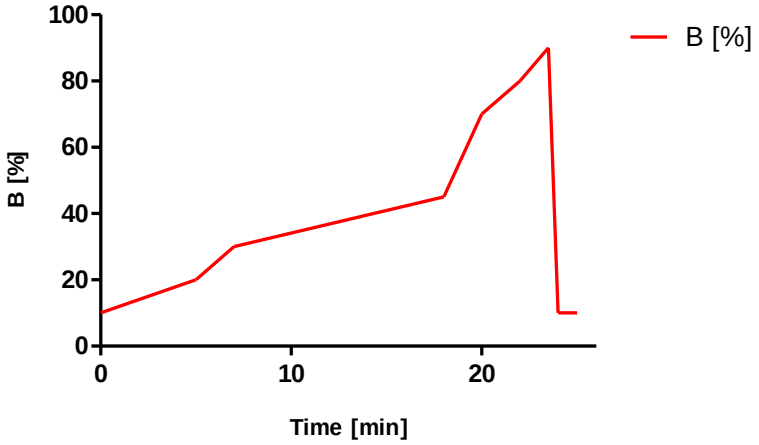


Figure 2.3: Graphical representation of the flow gradient used for Method I. B = 40% acetonitrile.

Table 2.3: Parameters including their settings used for Method I.

Parameter	Setting
Run time	25 min
Column	20 °C
Sampler	15 °C
UV	254 nm
Flow rate	0.5 mL/min
Loop	200 µL
Injection volume	50 µL
Injection type	Partial

As anticipated, injection of the DNA sequences alone onto the HPLC column resulted in the chromatograms shown in **Figure 2.4**. The oligonucleotides were detected at the following retention times (RT): NF-κB-1 RT 7.505 min, NF-κB-2 RT 7.468 min, EGR-1 RT 7.921 min, AP-1 RT 8.003 and STAT3 RT 8.533 min. Using the flow gradient shown in **Figure 2.3**, satisfactory results were achieved in terms of DNA retention times since all sequences eluted between 7 and 9 min. However, as visible on the individual chromatographic profiles (**Figure 2.4**), the UV intensities of the peaks were very high (between ca. 1600 mAU and ca. 5000 mAU) suggesting that either high concentrations or high volumes of oligonucleotides were injected.

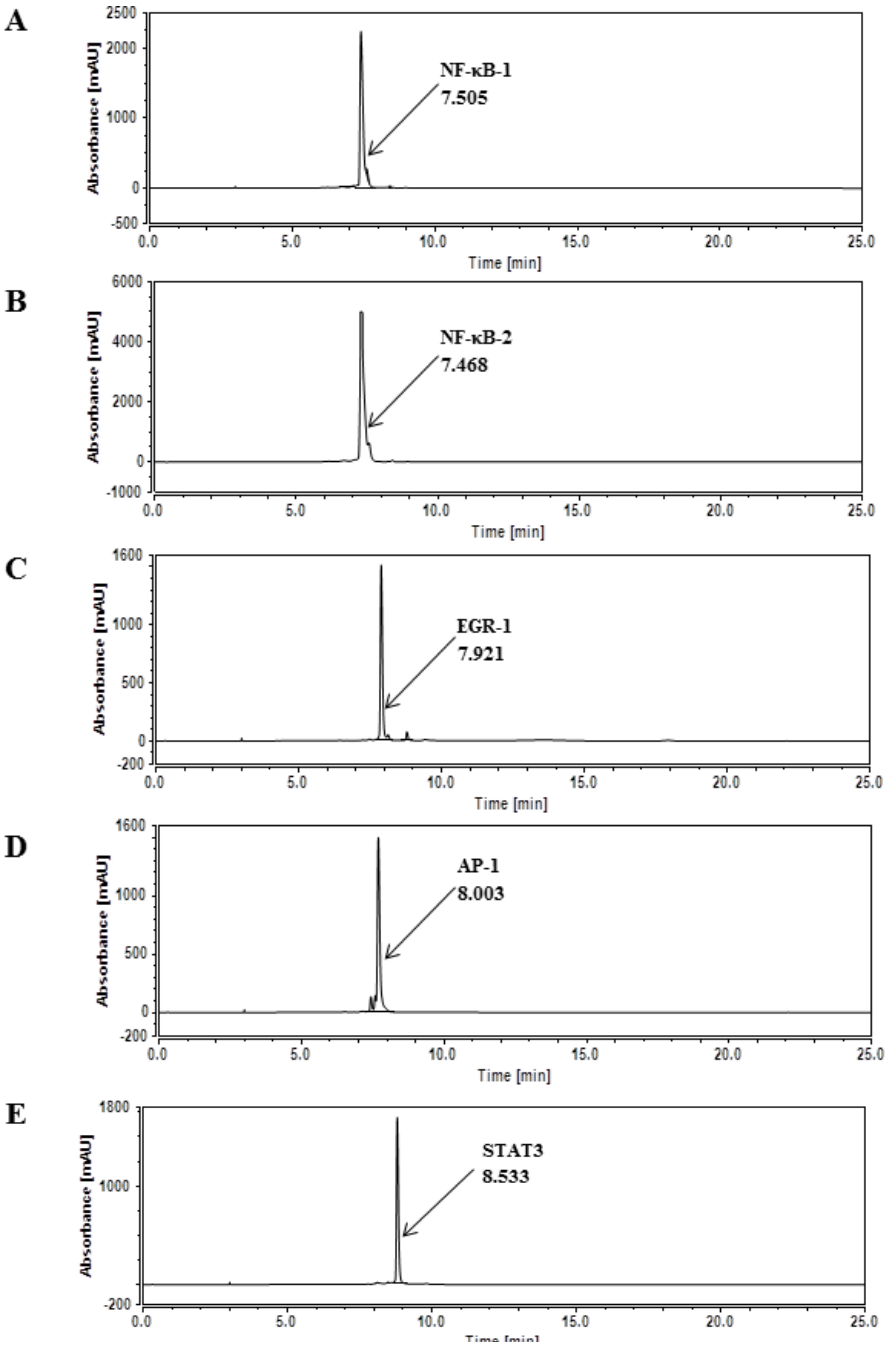


Figure 2.4: HPLC chromatograms of annealed NF-κB-1 (**A**), NF-κB-2 (**B**), EGR-1 (**C**), AP-1 (**D**) and STAT3 (**E**) sequences using Method A.

To address the high UV absorbances which were observed for Method I, two possible solutions were discussed: i) to reduce the concentration of the oligonucleotides by diluting the working solutions; or ii) to reduce the injection volume.

Previous RP-HPLC studies carried out on PBDs and short oligonucleotides have demonstrated that best results were achieved when a 4:1 ratio of PBD/DNA was used[96, 136, 157]. Therefore, the sample injection volume was reduced in order to decrease the UV signal of the DNA.

Method Ia

To reduce the sample injection volume, the previous 200 µL injection loop was changed to a 20 µL loop. The injected sample volume was reduced from 50 µL to 20 µL. The flow gradient and instrument settings were kept as shown in **Table 2.2** and **Figure 2.3**. The instrument settings are listed in **Table 2.4**. The hairpin DNA was then subjected to HPLC analysis using this method. The chromatographic profiles are shown in **Figure 2.5**. The hairpin oligonucleotides were detected at the following retention times: NF-κB-1 RT 7.318 min, NF-κB-2 RT 7.279 min, EGR-1 RT 7.654 min, AP-1 RT 7.837 and STAT3 RT 8.339 min. The loop change and the reduction of the injection volume slightly shifted the DNA retention times. Similar to Method I, the retention times were satisfactory (between 7 min and 9 min), however, the loop change and the reduced injection volumes resulted in a significant decrease in the UV absorbance of each oligonucleotide (from ca. 12.5 mAU to ca. 70 mAU) (**Figure 2.5**), making a precise detection and subsequent quantification difficult.

Table 2.4: Parameters including their settings used for optimised Method Ia.

Parameter	Setting
Run time	25 min
Column	20 °C
Sampler	15 °C
UV	254 nm
Flow rate	0.5 mL/min
Loop	20 µL
Injection volume	20 µL
Injection type	Partial

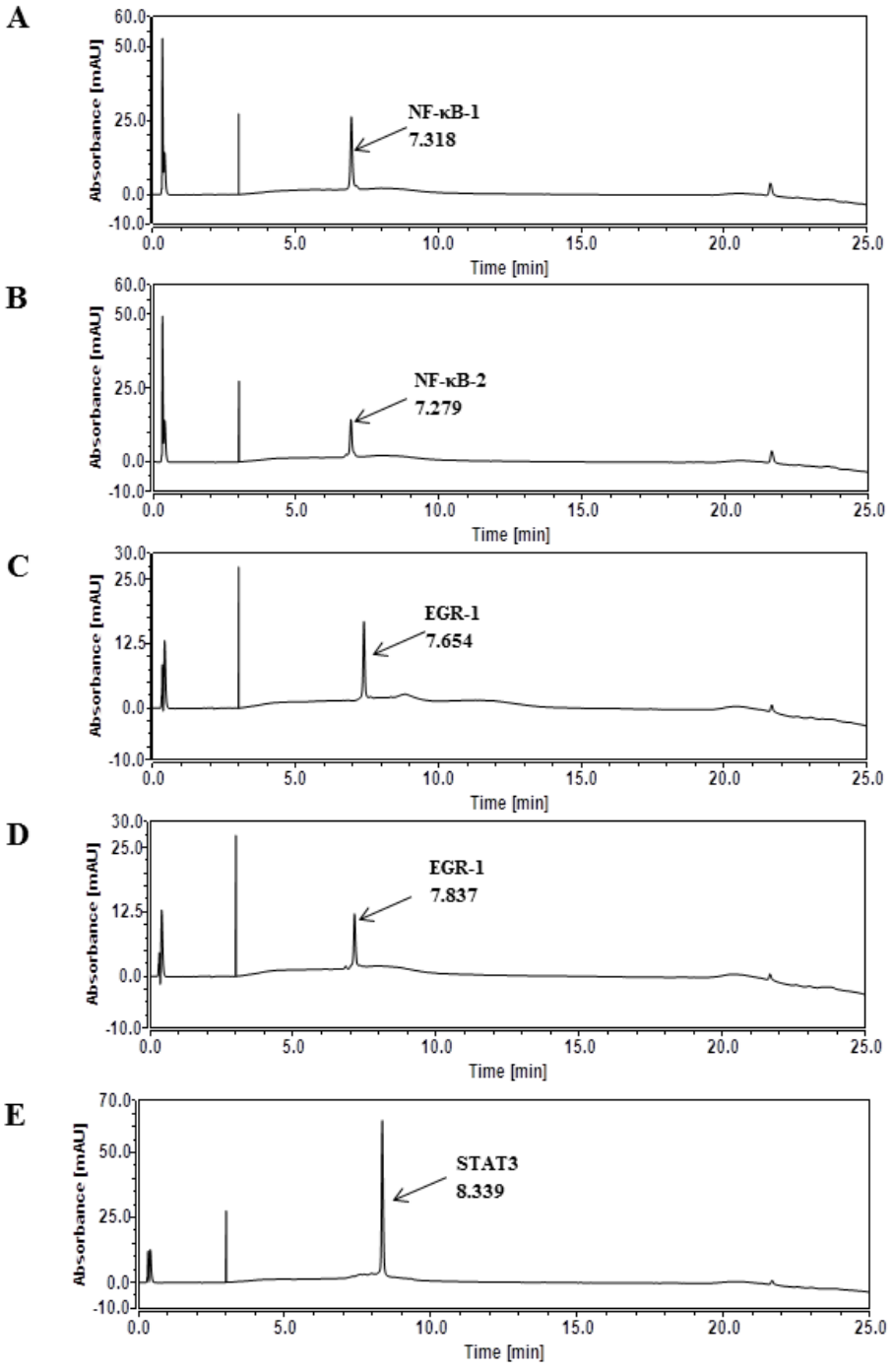


Figure 2.5: HPLC chromatograms of annealed NF-κB-1 (**A**), NF-κB-2 (**B**), EGR-1 (**C**), AP-1 (**D**) and STAT3 (**E**) sequences using Method Ia.

Method Ib

To overcome the problem described for Method Ia, the 20 µL injection loop was changed back to the previous 200 µL loop and the injection volume was increased to 30 µL. The instrument settings used for Method Ib are described in **Table 2.5**. The flow gradient applied for Method Ib is listed in **Table 2.2**. Similar to the procedure carried out for Method I and Ia, the hairpin oligonucleotides were subjected to RP-HPLC analysis. The obtained chromatographic profiles are shown in **Figure 2.6**. The DNA hairpins were observed at the following retention times: NF-κB-1 RT 7.456 min, NF-κB-2 RT 7.468 min, EGR-1 RT 7.921 min, AP-1 RT 8.026 and STAT3 RT 8.548 min. The loop change and the increase of the injection volume had an effect on the DNA elution times. The obtained retention times (between 7 and 9 min) and the UV absorbance (ca. 1000 mAU to ca. 1800 mAU) (**Figure 2.6**) for the hairpin oligonucleotides were considered as being satisfactory. Following this, Method Ib was used to investigate the interaction of GWL-78 and the NF-κB-1, NF-κB-2, EGR-1, AP-1 and STAT3 sequences (**Figure 2.1**) as described in the following Section 2.4.1.1.

Table 2.5: Parameters including their settings used for Method Ib.

Parameter	Setting
Run time	25 min
Column	20 °C
Sampler	15 °C
UV	254 nm
Flow rate	0.5 mL/min
Loop	200 µL
Injection volume	30 µL
Injection type	Partial

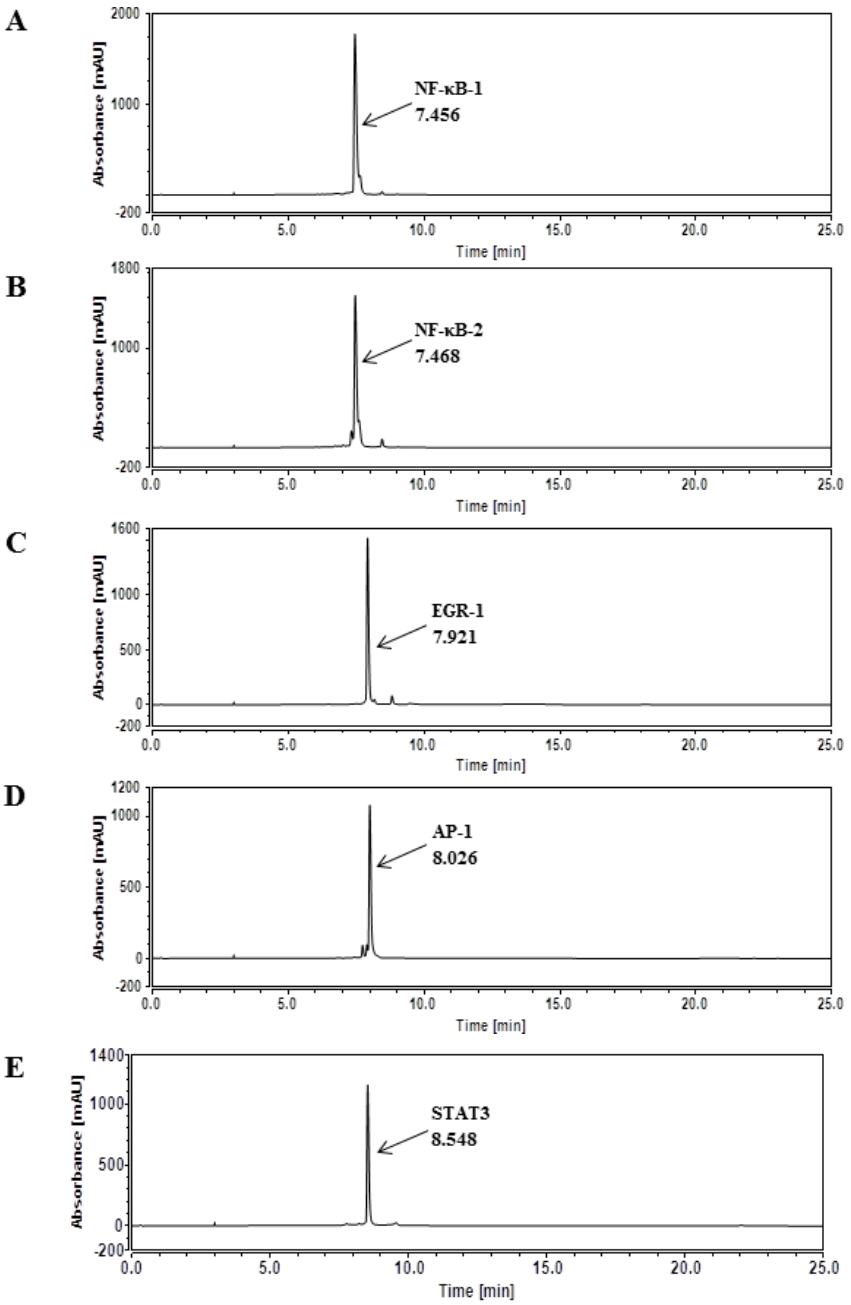


Figure 2.6: HPLC chromatograms of annealed NF-κB-1 (**A**), NF-κB-2 (**B**), EGR-1 (**C**), AP-1 (**D**) and STAT3 (**E**) sequences using Method Ib.

2.4.1.1 Interaction of GWL-78 with the cognate sequences of the oncogenic transcription factors NF-κB, EGR-1, AP-1 and STAT3 using Method Ib

The preparation procedure of hairpin DNA working solutions is described in detail in Section 7.2.1.2. Working solutions of GWL-78 were prepared according to the protocol described in Section 7.2.1.7. GWL-78/DNA complexes were prepared as described in Section 7.2.1.10. Briefly, a GWL-78 working solution of 100 μM was added to a DNA working solution of 25 μM in a 4:1 ratio (GWL-78/DNA). The mixture was mixed and incubated for various time intervals at 25 °C. A time-course study was carried out from 5 min to 24 hours by withdrawing samples at particular time points and immediate subjection to RP-HPLC analysis. The amount of adduct formed was calculated by measuring the Area Under the Curve (AUC) of the chromatographic peaks and plotted against the time.

The results obtained from the time-course experiment conducted on GWL-78 and NF-κB-1, NF-κB-2, EGR-1, AP-1 and STAT3 sequences are summarised in **Table 2.6** and **Figure 2.7**. For NF-κB, two possible consensus sequences were studied referred to as NF-κB-1 and NF-κB-2. It was expected that longer incubation time would result in greater adduct formation unless the reaction has reached saturation and the % adduct formed would remain at a constant value. However, significant variations in the extent of adduct formation were observed for all the sequences. This inconsistency in adduct formation was more pronounced for the EGR-1 sequence (**Table 2.6**, **Figure 2.7**). For example, 1.4% adduct were formed after 5 min incubation which increased to 47.5% after 15 min. However, the amount of adduct formed decreased to 37.4% after the 30 min incubation and increased again to 47.4% after the 3 hours incubation. More drastically, only 8% of adducts were formed after the 6 hours incubation which increased again to 35.7% after 24 hours.

Table 2.6: Extent of adduct formation (in %) during the time-course study (5 min to 24 hours) carried out on the interaction of GWL-78 and NF-κB-1, NF-κB-2, EGR-1, AP-1 and STAT3 sequences using Method Ib.

Time [min]	NF-κB-1	NF-κB-2	EGR-1	AP-1	STAT3
0	0.0	0.0	0.0	0.0	0.0
5	18.7	1.8	1.4	100.0	25.5
10	18.0	2.6	23.6	99.8	44.5
15	19.2	39.6	47.5	98.2	46.4
30	16.6	30.6	37.4	73.0	45.7
60	37.3	43.7	40.1	80.5	55.0
120	33.7	53.4	33.4	80.2	54.9
180	34.2	53.0	47.4	77.3	45.6
360	45.3	62.0	8.0	81.5	45.1
720	35.0	61.3	28.5	87.0	46.7
1440	51.8	69.7	35.7	79.7	51.3

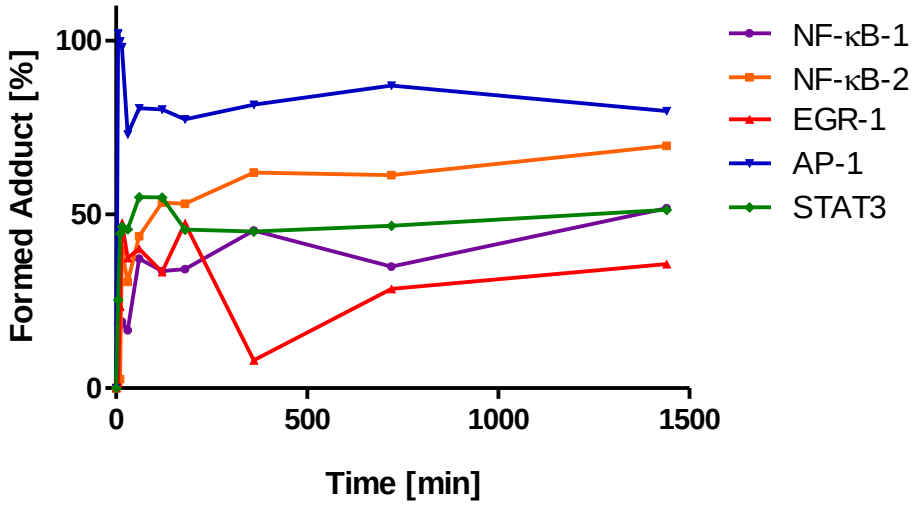


Figure 2.7: Graphical representation showing variability in adduct formation between GWL-78 and the sequences NF-κB-1, NF-κB-2, EGR-1, AP-1 and STAT3 using Method Ib.

These differences in the extent of adduct formation suggest a loss of sample during the injection process within the sample loop. The question was raised whether the observed discrepancies in the extent of adduct formation may be due to the partial-loop injection mode used for Method Ib.

The sample loop is used to precisely measure the volume of sample to be injected. During HPLC two different injection modes can be applied to deliver the sample to the column, the partial-loop and the full-loop injection mode. The injection volume necessary for the partial-loop injection is less than the loop volume and should be ideally less than half the loop volume. The autosampler withdraws the injection volume plus 2x the cut volume which is specified and dependent on the desired sample volume to be injected (usually 1-30 µL). The cut volume is positioned before and after the sample but is not injected as it is "cut" from the sample by the injection valve. The sample needle aspirates the specified injection volume plus the cut volume and transfers it to the sample loop. Following this, the initial mobile phase, sample and the wash solvent which is necessary for the needle wash are co-injected onto the HPLC column. This co-injection may lead to sample dilution and loss. The full-loop injection mode is more accurate and precise than the partial-loop injection mode as the injection volume is equal to the loop volume and the sample manager injects 100% of the sample loop volume. To flush the sample loop, the autosampler withdraws, at least, four times the volume of the sample loop from the vial and injects the centre of the larger sample volume. Full-loop injections deliver the best results in terms of reproducibility, carryover and linearity but require higher sample consumption.

Method Ic

To investigate whether the inconsistency in the extent of adduct formation was due the injection mode, the previous 200 µL loop was changed to a 50 µL loop applying a full-loop injection of 50 µL sample. The loop change was carried out to reduce the sample volume that is being injected as the full-loop injection mode for the 200 µL loop would imply a sample injection volume of 200 µL. The instrument settings used for Method Ic are listed in **Table 2.7**. The flow gradient used for Method Ic is shown in **Table 2.2** and **Figure 2.3**. Following this, the hairpin oligonucleotides were subjected to RP-HPLC analysis according to the procedure described for Method I (see also Section 2.4.1). The

obtained chromatographic profiles are shown in **Figure 2.8**. The DNA hairpins were detected at the following retention times: NF-κB-1 RT 7.041 min, NF-κB-2 RT 6.948 min, EGR-1 RT 7.450 min, AP-1 RT 7.156 and STAT3 RT 8.491 min. Similarly, the loop and injection mode change affected the DNA elution times compared to Method Ib. Using Method Ic, satisfactory results were achieved in terms of DNA retention times (7 to 9 min) and UV absorbance (ca. 1000 mAU to ca. 1800 mAU) (**Figure 2.8**). Method Ic was then used to evaluate the interaction of GWL-78 and the NF-κB-1, NF-κB-2, EGR-1, AP-1 and STAT3 sequences (**Figure 2.1**) as described in the following Section 2.4.1.2.

Table 2.7: Parameters including their settings used for Method Ic.

Parameter	**Setting**
Run time	25 min
Column	20 °C
Sampler	15 °C
UV	254 nm
Flow rate	0.5 mL/min
Loop	50 µL
Injection volume	50 µL
Injection type	full loop

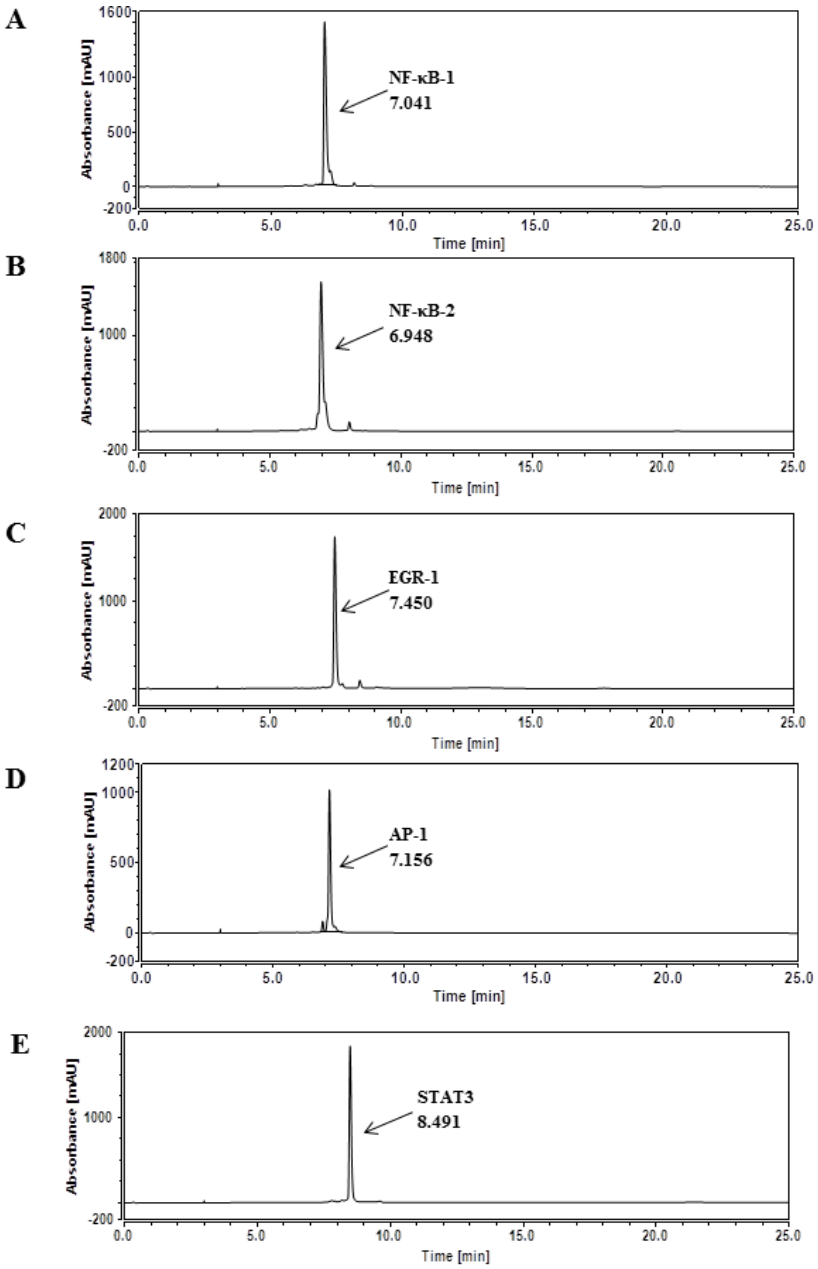


Figure 2.8: HPLC chromatograms of annealed NF-κB-1 (**A**), NF-κB-2 (**B**), EGR-1 (**C**), AP-1 (**D**) and STAT3 (**E**) sequences using Method Ic.

2.4.1.2 Interaction of GWL-78 with the cognate sequences of the oncogenic Transcription factors NF-κB, EGR-1, AP-1 and STAT3 using Method Ic

The preparation procedure of hairpin DNA working solutions is described in Section 7.2.1.2 in detail. Working solutions of GWL-78 were prepared according to the protocol described in Section 7.2.1.7. GWL-78/DNA complexes were prepared according to the procedure described in Section 7.2.1.10. The results obtained from the RP-HPLC study are summarised in **Tables 2.8**, **2.9**, **2.10** and **Figure 2.9**. The time-course experiment showed no variations in the extent of reaction when Method Ic was applied (**Figure 2.9**). This result confirmed that the full-loop injection produced better results in terms of reproducibility and linearity compared to the partial-loop injection mode as discussed in Section 2.4.1.1 in detail.

Initially, using Method Ic, each hairpin oligonucleotide (25 μM) was injected three times onto the HPLC in order to obtain the experimental error and investigate whether Method 1c provides consistent retention times for each hairpin oligonucleotide. **Table 2.8** shows the obtained RT for each oligonucleotide after three injections and the average RT in correspondence with their standard deviation (SD). The corresponding HPLC chromatograms are shown in the Appendix (**Figure A2.1 Appendix**). Based on these retention times, new peaks that appeared in the chromatographic profiles after incubation with SJG-136 and were also not within the SD were assumed to be SJG-136/DNA adducts.

Table 2.8: Obtained retention times for the DNA sequences NF-κB-1, NF-κB-2, EGR-1, AP-1 and STAT3.

Sequence	RT [min]	Average RT [min] ± SD [min]
NF-κB-1	7.178, 7.041, 7.006	7.075 ± 0.091
NF-κB-2	7.198, 7.072, 6.904	7.058 ± 0.147
EGR-1	7.673, 7.369, 7.402	7.481 ± 0.167
AP-1	7.267, 7.278, 7.325	7.290 ± 0.031
STAT3	8.535, 8.474, 8.328	8.446 ± 0.106

The data from the RP-HPLC experiments have shown that GWL-78 may interact with the cognate sequences of the oncogenic transcription factors NF-κB, EGR-1, AP-1 and STAT3 as adduct formation was observed for all studied sequences (**Table 2.9**, **Table 2.10** and **Figure 2.9**). The differences in RT between some of the DNA sequences (*e.g.*, NF-κB-1, NF-κB-2 and AP-1) and the corresponding adducts are not statistically significant. However, subsequent MALDI-TOF analysis (after 5 min, 3 hours and 24 hours incubation, all MALDI-TOF spectra included in the Appendix) has clearly confirmed that adducts had been formed. The study showed differences in extent and rate of reaction between GWL-78 and the DNA hairpins. For example, GWL-78 reacted rapidly with sequences NF-κB-1 (~42% adduct formed) (**Figure A2.2B Appendix**), AP-1 (~100% adduct formed) (**Figure A2.8B Appendix**) and STAT3 (~42% adduct formed) (**Figure A2.10B Appendix**) after 5 min. However, as the incubation proceeded, GWL-78 showed greater reactivity with the NF-κB-2 (~100% adduct formed after 3 hours) (**Figure A2.4C Appendix**) and AP-1 (~100% adduct formed after 3 hours) (**Figure A2.8C Appendix**) sequences whereas moderate adduct formation was observed for the NF-κB-1 (~51% adduct formed after 3 hours (**Figure A2.2C Appendix**) , ~77% adduct formed after 24 hours (**Figure A2.2D Appendix**)), EGR-1 (~10% adduct formed after 3 hours (**Figure A2.6C Appendix**), ~32% adduct formed after 24 hours (**Figure A2.6D Appendix**),) and STAT3 (~43% adduct formed after 3 hours (**Figure A2.10C Appendix**), ~49% adduct formed (**Figure A2.10D Appendix**)) sequences (**Table 2.9 and Figure 2.9**). Moreover, GWL-78 formed one type of adduct with NF-κB-2 and AP-1 sequences. Multiple adducts were observed for NF-κB-1, EGR-1 and STAT3 sequences (**Figures A2.2**, **A2.4**, **A2.6**, **A2.8**, **A2.10** and **Table 2.10**). These observations suggest a preference of GWL-78 toward the NF-κB-1 and AP-1 sequences. However, the differences in the RTs between NF-κB-1, NF-κB-2 and AP-1 sequences and the corresponding adducts were very small (NF-κB-1 RT 7.069 min, GWL-78/NF-κB-1 adduct mean RT 7.296 min; NF-κB-2 RT 7.097 min, GWL-78/NF-κB-2 adduct mean RT 7.219 min; AP-1 RT 7.206 min, GWL-78/AP-1 adduct mean RT 7.477 min) (**Figures A2.2**, **A2.4** and **A2.8 Appendix**). This was assumed to be due to the high GC content of the hairpin oligonucleotides used in this study. GC-rich (more polar than AT-rich) sequences are very similar in polarity as their corresponding PBD/DNA adducts. As a consequence, the parent DNA sequences and the corresponding adducts elute at very close retention times. These observations were in broad agreement with previous studies where similar small differences (<1 min) between GC-rich DNA sequences and their corresponding

PBD/DNA adducts have been reported[96, 200]. In contrast to this, the retention time differences between AT-rich DNA sequences (less polar than GC-rich) and their corresponding adducts are bigger as the non-polar character of the PBD/DNA complex is increased after drug binding[136]. This was also the case in the current thesis when the reactive guanines within sequence AP-1 were consecutively replaced with the less polar inosine bases during the DNA-binding site analysis on the AP-1 sequence (see Chapter 5). During this study it was observed that sequences bearing less guanine bases have significant differences in the retention times compared to the corresponding adducts[201]. For example, when the parent AP-1 hairpin sequence has two guanine bases replaced with two inosine bases (AP-1 1, 17I, **Figure 5.3**), the corresponding adducts show much bigger differences in the retention times (adduct 1 RT 12.50 min; adduct 2 RT 14.38 min) compared to the AP-1 1, 17I sequence alone (RT 6.50 min) (**Figure 5.4**). Furthermore, it was not conclusive whether the observed adduct peaks were only the adducts or a mixture of the parent DNA and the corresponding adducts. Limitations of the developed technique may also be an explanation for this observation. It is possible that multiple adducts had been formed but using the developed methodology the parent DNA and the adduct peaks could not accordingly be resolved. Therefore, whether the reaction between NF-κB-2 and AP-1 sequences and GWL-78 proceed to full extent with 100% adduct formed after 5 min cannot be confirmed by utilising this method. Optimisation of the developed method has to be carried out in future studies in order to resolve those peaks. In summary, although no significant differences in the retention times of some DNA and their corresponding adducts have been observed, an indication that no adduct had been formed, adduct formation was positively confirmed by MALDI-TOF-MS for all studied sequences. Due to space limitations, as an example, only the chromatographic profiles obtained for the interaction of GWL-78 with the NF-κB-2 sequence using Method Ic are shown below (**Figure 2.10**). All obtained HPLC chromatograms are shown in the **Appendix**. In the chromatographic profiles for the interaction between GWL-78 and the EGR-1, AP-1 and STAT3 sequences, the peak area for GWL-78 is present for all studied sequences. However, they also show variabilities in the peak associated to the compound (**Figures A2.6**, **A2.8** and **A2.10 Appendix**). As the reaction proceeds and more adduct is formed, the peak area corresponding to the compound is expected to decrease. The HPLC data show that in the cases for EGR-1, AP-1 and STAT3 the peak area for GWL-78 increases, although more adduct is formed after 3 hours and 24 hours, respectively (**Figures A2.6**, **A2.8** and **A2.10 Appendix**). This can be explained through the overall

detected absorbance. For example, in **Figure A2.6B** (**Appendix**) the maximum absorbance measured is ~1600 mAU (after 5 min incubation of GWL-78 with EGR-1 sequence. However, in **Figure A2.6D** (**Appendix**) the maximum detected absorbance is only ~500 mAU (after 24 hours incubation). Hence, the peak associated to GWL-78 in **Figure A2.6D** appears much higher than in **Figure A2.6D** despite more adduct being formed. For example, the calculated AUC for GWL-78 in the HPLC chromatogram **Figure A2.6B**, **C** and **D** (**Appendix**) are 101.15 mAU*min (after 5 min incubation), 80.69 mAU*min (3 hours incubation) and 72.35 mAU*min (after 24 hours). This clearly demonstrated that more adduct had been formed and the drug was used up despite this not being visible on the chromatographic profiles. If the chromatographic profiles would be presented in the same way (fixed maximum absorbance) then the peak corresponding to the compound in **Figure A2.6D** would appear much smaller. These data were also supported by the literature where the same observation was made for PBD monomers and dimers[96, 136].

Table 2.9: The extent of adduct formation between GWL-78 and the TF sequences used in this study (adduct formed in %) at different time points.

Time [min]	NF-κB-1	NF-κB-2	EGR-1	AP-1	STAT3
5	42	6	0	100	42
180	51	100	10	100	43
1440	77	100	32	100	49

Table 2.10: The number of adducts formed between GWL-78 and the TF sequences used in this study (adduct formed in %) after 24 hours of incubation (**Figures A2.2**, **A2.4**, **A2.6**, **A2.8**, **A2.10 Appendix**).

	NF-κB-1	NF-κB-2	EGR-1	AP-1	STAT3
GWL-78	2	1	3	1	3

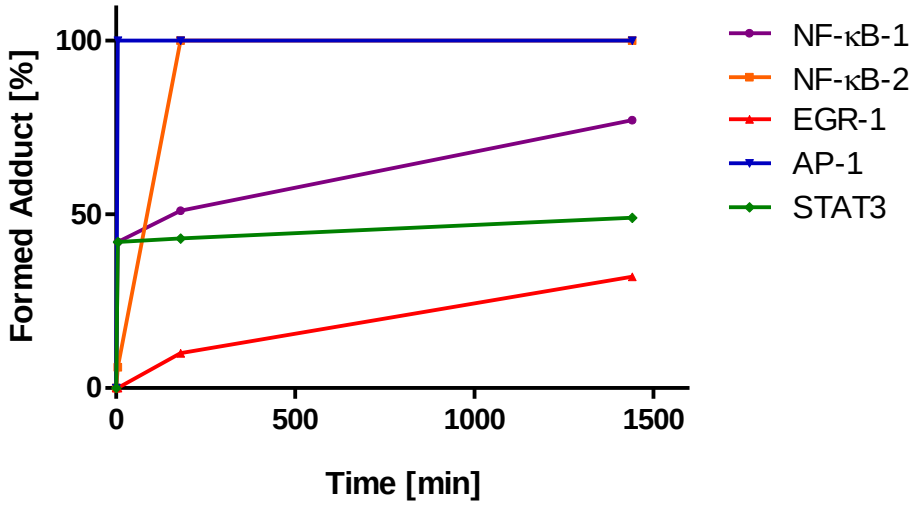


Figure 2.9: Graphical representation showing variability in adduct formation after 5 min, 3 hours and 24 hours.

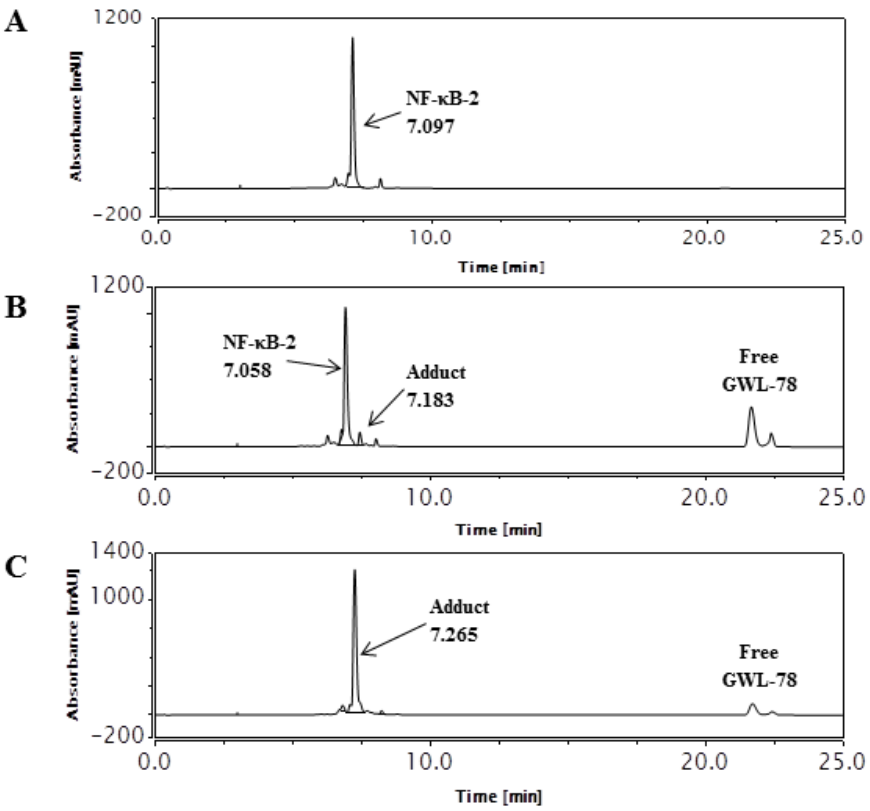


Figure 2.10: Time-course HPLC chromatograms; (**A**) Annealed NF-κB-2 sequence; (**B**) Annealed NF-κB-2 sequence after incubating with GWL-78 for 5 min showing the appearance of one new minor peak at RT 7.183 min; (**C**) Annealed NF-κB-2 sequence after incubating with GWL-78 for 24 hours showing conversion to one main adduct at RT 7.265 min.

2.4.1.3 Interaction of KMR-28-39 with the cognate sequences of the oncogenic transcription factors NF-κB, EGR-1, AP-1 and STAT3 using Method Ic

Method Ic was then applied to study the interaction of the PBD monomer KMR-28-39 and the hairpin sequences shown in **Figure 2.1**. The preparation procedure of hairpin DNA working solutions is described in Section 7.2.1.2. Working solutions of KMR-28-39 were prepared as described in Section 7.2.1.8. KMR-28-39/DNA complexes were prepared according to the procedure described in Section 7.2.1.10. The results obtained from the RP-HPLC study are summarised in **Table 2.11** and **Table 2.12**. Similar to the results for GWL-78, the time-course study showed no fluctuations in the extent of reaction when Method Ic was applied, again confirming that the full-loop injection mode yields reproducible and linear results compared to partial-loop injection.

The results from the RP-HPLC study have shown that KMR-28-39 may also be able to bind to the cognate sequences of the oncogenic transcription factors NF-κB, EGR-1, AP-1 and STAT3. Adduct formation was observed for all studied sequences. In contrast to the study conducted on GWL-78, a similar extent of reaction was observed for all sequences (**Table 2.11** and **Table 2.12**). According to the obtained chromatographic profiles (**Figures A2.12**, **A2.14**, **A2.16**, **A2.18** and **A2.20 Appendix**), KMR-28-39 reacted the full extent (100% adduct formation already after 5 min incubation (**Figures A2.12B**, **A2.14B**, **A2.16B**, **A2.18B** and **A2.20B Appendix**)) and formed one main adduct with the NF-κB-1, NF-κB-2, AP-1 and STAT3 sequences (**Figures A2.12B**, **A2.14B**, **A2.18B** and **A2.20B Appendix**). One major and one minor adduct was observed for the EGR-1 sequence (**Figure A2.16B Appendix**). Based on these data, KMR-28-39 demonstrated to be highly selective for the hairpin sequences used in this study. However, these results do not confirm that KMR-28-39 immediately (after 5 min) formed one adduct with the NF-κB-1, NF-κB-2, AP-1 and STAT3 sequences, and one major and one minor adduct with the EGR-1 sequence as these observations were assumed to be a consequence of the insufficient resolution of the DNA and KMR-28-39/DNA peaks, resulting in a simultaneous elution of the DNA and ligand/DNA complexes from the column. Therefore, KMR-28-39 may form multiple adducts with the hairpin sequences used in this study which elute from the HPLC column at the same time due to similar polarity of the DNA and KMR-28-39/DNA adducts. This results in the appearance of one major peak and is not conclusive enough to state that KMR-28-39 has reacted to full extent with those

sequences after 5 min. With this consideration in mind, the HPLC data from the interaction of KMR-28-39 and the NF-κB-1, NF-κB-2, EGR-1, AP-1 and STAT3 need to be analysed with caution. As an example, only the chromatographic profiles obtained for the interaction of KMR-28-39 with the NF-κB-1 sequence using Method Ic are shown below (**Figure 2.11**). All obtained HPLC chromatograms are shown in the **Appendix**. Additionally, all chromatographic profiles (except for EGR-1) lack the appearance of the peak associated to KMR-28-39 after 5 min, 3 hours and 24 hours (**Figures A2.12**, **A2.14**, **A2.18** and **A2.20 Appendix**). This was explained through the previously reported high reactivity of KMR-28-39 towards GC-rich DNA sequences[90]. Based on these data, it was assumed that KMR-28-39 reacted rapidly. This observation was investigated using MALDI-TOF mass spectrometry method. The results from this study have positively confirmed adduct formation after 5 min, 3 hours and 24 hours as the correct masses for the 1:1 KMR-28-39/DNA adducts were detected (see Section 2.4.2.1). As discussed in Section 2.4.1.2 in detail, the lack of the KMR-28-39 peak may also be due to the high detected absorbance of minimum 1000 mAU, which affects the intensity of the KMR-28-39 peaks.

Table 2.11: The extent of adduct formation between KMR-28-39 and the TF sequences used in this study (adduct formed in %) after a 24 hours incubation.

	NF-κB-1	**NF-κB-2**	**EGR-1**	**AP-1**	**STAT3**
KMR-28-39	100	100	100	100	100

Table 2.12: The number of adducts formed between KMR-28-39 and the TF sequences used in this study (adduct formed in %) after 24 hours of incubation (**Figures A2.12**, **A2.14**, **A2.16**, **A2.18** and **A2.20 Appendix**).

	NF-κB-1	**NF-κB-2**	**EGR-1**	**AP-1**	**STAT3**
KMR-28-39	1	1	2	1	1

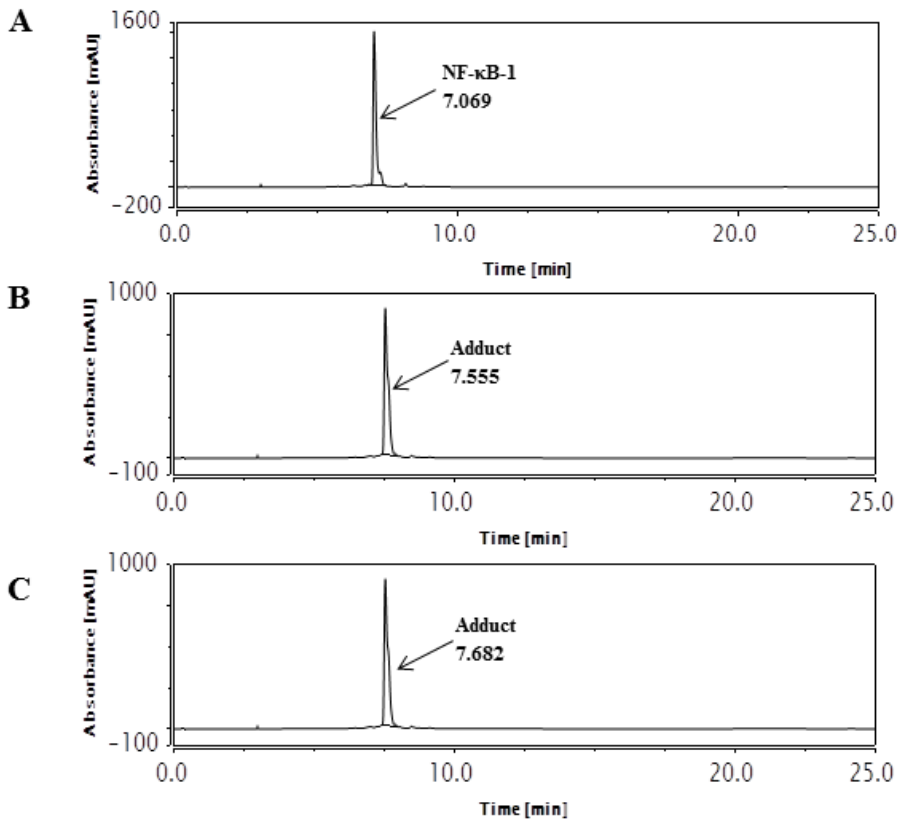


Figure 2.11: Time-course HPLC chromatograms; (**A**) Annealed NF-κB-1 sequence; (**B**) Annealed NF-κB-1 sequence after incubating with KMR-28-39 for 5 min showing the appearance of one new major peak at RT 7.555 min; (**C**) Annealed NF-κB-1 sequence after incubating with KMR-28-39 for 24 hours showing no further changes in the chromatographic profile with one major peak at 7.682.

2.4.2 MALDI-TOF-MS method development

After completing the RP-HPLC experiments using the PBD monomers GWL-78 and KMR-28-39 applying Method Ic, the goal was to develop a MALDI-TOF-MS method to confirm the stoichiometry of the adducts formed during the HPLC study.

The MALDI-TOF experiments were carried out on a Bruker Daltonics Autoflex™ (Bruker Daltonik) with an automated high-throughput Matrix-Assisted Laser Desorption/Ionisation Time of Flight (MALDI-TOF) system. A nitrogen laser in positive linear mode with a delayed extraction of 500 ns and an accelerating voltage of 25000 V was applied during this study. Acquisition was undertaken between 1000-10000 Dalton with 100 shots per spectrum. The instrument was calibrated prior to sample analysis using insulin as standard. Briefly, 1 µL insulin was spotted onto the MALDI-TOF target plate and naturally dried, followed by addition of 1 µL α-Cyano-4-hydroxycinnamic acid (CHCA) matrix. All data were processed using AutoFlex software (Bruker Daltonik).

Working solutions of the hairpin oligonucleotides, ligands and ligand/DNA complexes were prepared according to the procedure described in Section 7.2.3.2. After incubation of the ligands with the DNA, the samples were diluted in 1:1, 1:4 and 1:10 ratios with 0.1 M triethylammonium acetate buffer (TEAA). The ZipTip$_{C18}$™ reversed phase sample preparation method shown **Figure 2.12** was applied. The aim of this procedure was to purify and concentrate the oligonucleotide samples (femtomoles to picomoles) prior to MALDI-TOF analysis which yields better data quality. The samples were diluted with TEAA due to optimal binding and recovery of the DNA from C_{18} is performed in the presence of this ion-pairing agent.

In the first step, the ZipTip$_{C18}$™ was equilibrated for sample binding by aspirating and dispensing the equilibration solution composed of 50% acetonitrile and 0.1 M TEAA in a 1:1 ratio, followed by subsequent washing with a wash solution of 0.1 M TEAA. The equilibration and wash steps were repeated 3 to 4 times. The binding and washing of the oligonucleotides were carried out by aspirating the same ZipTip$_{C18}$™ into the sample solution. In order to achieve maximum binding, this step was repeated 10 times. Following this, the ZipTip$_{C18}$™ was washed for 3 cycles by aspirating and dispensing of the wash solution (0.1 M TEAA) and HPLC grade water for 3 cycles. The matrix used

for MALDI-TOF analysis was prepared according to the procedure described in Section 7.2.3.1. The ligand/DNA complexes were eluted from the ZipTip$_{C18}$™ and mixed with 0.5 µL of the matrix. 0.2 µL of sample/matrix mix was spotted onto the MALDI-TOF target plate (MTP AnchorChip 384TF, Bruker Daltonik) and naturally allowed to dry prior to MALDI-TOF-MS analysis.

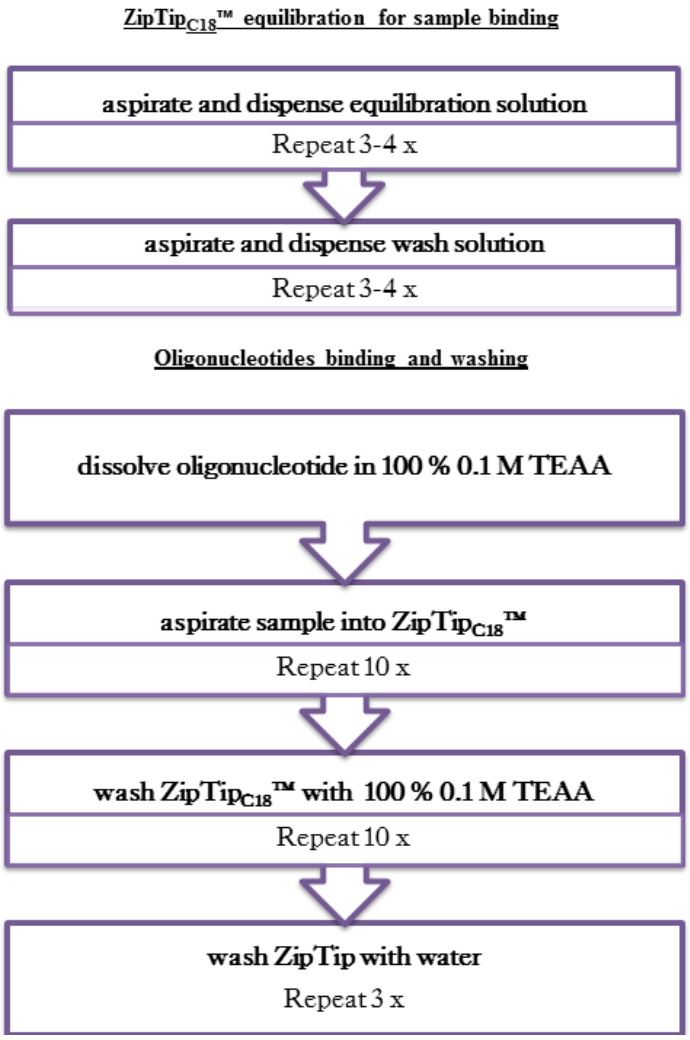


Figure 2.12: MALDI-TOF sample preparation flowchart using the ZipTip$_{C18}$™ method.

2.4.2.1 MALDI-TOF-MS study of GWL-78 and KMR-28-39 with the cognate sequences of oncogenic transcription factors NF-κB-1, EGR-1, AP-1 and STAT3

The MALDI-TOF method (described in Section 2.4.2) was then used to confirm the adduct formation between GWL-78 and KMR-28-39 and the NF-κB-1, NF-κB-2, EGR-1, AP-1 and STAT3, which was observed during the RP-HPLC study. The MALDI-TOF results for GWL-78 and KMR-28-39 are summarised in **Table 2.13**. All MALDI-TOF spectra are shown in the **Appendix**. The MALDI-TOF spectra obtained showed for both ligands and all studied sequences the correct molecular weight for the 1:1 ligand/DNA adducts (**Table 2.13**). For all studied DNA sequences adduct formation was observed after 5 min, 3 hours and 24 hours incubation with both ligands, GWL-78 (**Figures A2.3**, **A2.5**, **A2.7**, **A2.9** and **A2.11**) and KMR-28-39 (**Figures A2.13**, **A2.15**, **A2.17**, **A2.19** and **A2.21**). Based on the detected masses, the adducts formed between the ligands and the hairpin oligonucleotide were assumed to be the 1:1 ligand/DNA adducts, since the binding of the ligands to one guanine prevents further ligand association to the DNA sequence. Two examples of the obtained MS spectra are shown in **Figure 2.13**. Based on the chromatographic profiles it was anticipated that the SJG-136/DNA adduct peaks would have higher intensities compared to the parent DNA as more adduct is formed while the reaction proceeded. Interestingly, the MALDI-TOF results contrasted to this anticipation as higher intensity peaks were observed for the parent hairpins whereas the SJG-136/adduct peaks were of lower intensity. This indicated the presence of higher DNA content in the sample than the formed SJG-136/DNA adduct. Previous studies carried out on the interaction of SJG-136 with short oligonucleotides have shown that during the MALDI-TOF process the formed adduct separates into the parent DNA and SJG-136 resulting in a higher peak for the DNA and a smaller peak for the formed SJG-136/DNA adduct[136, 157, 202]. **Figure 2.14** demonstrates that although the reaction between the 12 bp dsDNA and SJG-136 had nearly gone to completion in the HPLC study (~95.5% adduct formed after 24 hours), only a small peak for the 1:1 SJG-136/DNA adduct (m/z 7844.5) was detectable in the MALDI-TOF spectra[157]. This previous observation was in broad agreement with the MALDI-TOF results obtained in this study. Moreover, past studies have demonstrated that PBD molecules do not bind to single stranded DNA, as they require a minor groove environment for covalent attachment[125]. Consequently, it possible that not annealed, linear DNA is still present in the sample that is detected by MALDI-TOF, therefore contributing to the peak intensity of the detected DNA peak[157].

All of the aspects mentioned in the discussion above may enhance the MALDI-TOF signals for the parent DNA and therefore, explain the appearance of MALDI-TOF DNA peaks of high intensities compared to SJG-136/DNA formed adducts.

Table 2.13: Theoretical and observed masses [Da] of single-stranded hairpin oligonucleotides and their 1:1 adducts with GWL-78 and KMR-28-39.

DNA Sequence	DNA Mass	Theoretical Mass DNA/GWL-78 (1:1) (DNA Mass + 590.6)	Observed DNA/GWL-78 Adduct Mass (1:1)	Theoretical Mass DNA/KMR-28-39 (1:1) (DNA Mass + 667)	Observed DNA/KMR-28-39 Adduct Mass (1:1)
NF-κB-1	7032.6	7623.2	7624.0	7699.6	7698.4
NF-κB-2	7033.6	7624.2	7625.1	7700.6	7701.3
EGR-1	6416.2	7006.8	7007.4	7083.2	7082.5
AP-1	5793.8	6384.4	6383.9	6460.8	6461.8
STAT3	6412.2	7002.8	7001.9	7079.2	7078.5

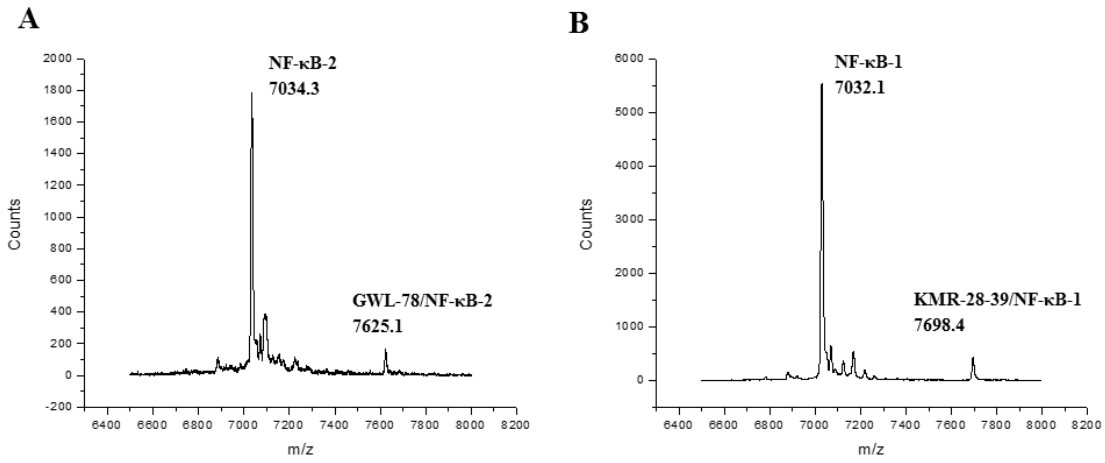


Figure 2.13: **A**, MALDI-TOF spectrum of GWL-78/NF-κB-2 (after a 24 hours incubation) confirming the stoichiometry of the 1:1 adduct formed (NF-κB-2 observed mass: 7034.3 m/z, theoretical mass: 7033.6 m/z, NF-κB-2 adduct observed mass: 7625.1 m/z, theoretical mass: 7624.2 m/z); **B**, MALDI-TOF spectrum of KMR-28-39/ NF-κB-1 (after a 24 hours incubation) confirming the stoichiometry of the 1:1 adduct formed (NF-κB-1 observed mass: 7032.1 m/z,

theoretical mass: 7033.6 m/z, NF-κB-1 adduct observed mass: 7698.4 m/z, theoretical mass: 7699.6 m/z).

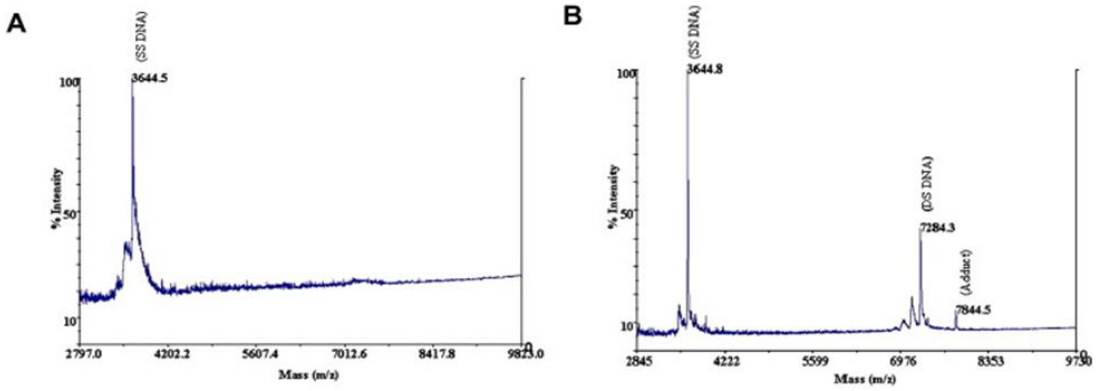


Figure 2.14: MALDI-TOF spectra confirming the identity of SJG-136 forming adduct with 12 bp long dsDNA. (**A**) MALDI–TOF MS spectrum of dsDNA alone showing that dsDNA separates into the corresponding single stranded DNA under the MALDI-TOF conditions. (**B**) MALDI–TOF MS spectrum of the 1:1 SJG-136/dsDNA adduct. The ion at *m*/*z* 7844.5 corresponds to the $[M+H]^+$ ion of the SJG-136/dsDNA adduct, and the ions at 7284.3 and 3644.8 correspond to $[M+H]^+$ ions of dsDNA and ssDNA, respectively, formed by in source fragmentation.

2.5 Summary of RP-HPLC/MS method development

In summary, after the development of a suitable RP-HPLC/MALDI-TOF method, it was applied to study the interaction of the PBD monomers GWL-78 and KMR-28-39 with the cognate sequences of the oncogenic transcription factors NF-κB, EGR-1, AP-1 and STAT3, which have recently been demonstrated to interact with consensus transcription factor sequences[90, 102]. The results obtained indicated that both ligands may bind to the NF-κB, EGR-1, AP-1 and STAT3 sequences. In all experiments adduct formation was observed during the RP-HPLC study and their 1:1 stoichiometry subsequently confirmed by MALDI-TOF-MS. These data validated the developed method and demonstrated that it may be suitable for its application in future studies. Following method validation, it was used to investigate the interaction of the PBD dimer SJG-136 with the same hairpin sequences. The results obtained from this study are presented in the following chapter (Chapter 3).

Chapter 3: Evaluation of the interaction of the PBD dimer SJG-136 with cognate sequences of oncogenic transcription factors by biophysical analysis

3.1 Objectives

The main aim of this study was to investigate the interaction of the PBD dimer SJG-136 with the cognate sequences of the oncogenic transcription factors NF-κB, EGR-1, AP-1 and STAT3 (**Figure 3.1**) using the HPLC/MS analytical method developed in Chapter 2. These particular consensus transcription factor sequences were selected for the study as they are known to play a crucial role in the development of human malignancies and to be over-expressed in human cancer types in which SJG-136 has demonstrated to be highly efficient[158]. For example, NF-κB is over-expressed in most leukaemias, STAT3 in breast cancer and EGR-1 and AP-1 in ovarian cancer. Extensive studies have been carried out to evaluate the cytotoxicity of SJG-136 on a panel of human tumour cell lines[158]. The agent exhibited LC_{50} (median lethal concentration) values ranging from 0.14 to 320 nmol/L[158]. The results clearly demonstrated selectivity of SJG-136 toward particular cell lines and could not be explained through the previously reported mechanism of action, including DNA strand breakage[94], inhibition of enzymes[101, 203] and arrest of the replication fork. Moreover, the NCI cell line screen has revealed an activity pattern for SJG-136 similar to some reported DNA-binding agents (*e.g.*, melphalan, cyclophosphamide, chlorambucil). However, its overall profile did not exactly match with cluster patterns associated with any specific chemotherapeutic agents, hence assuming a unique mechanism of action of SJG-136.

Therefore, new data relating to any potentially additional mechanism of action of SJG-136 could help to understand and explain the previously observed differences in activity between various human cancer types (*e.g.*, *in vitro* 60 Cell Line Screen[158]). To address this issue, in our study, SJG-136 was incubated with the hairpin forming oligonucleotides depicted in **Figure 3.1** at 25 ºC for various time intervals, followed by subjection to ion pair reversed phase HPLC analysis. The extent and rate of adduct formation were calculated by measuring the area under the curve (AUC). The stoichiometry of adducts formed was confirmed by MALDI-TOF-MS analysis based on the molecular weight. The

fluorescence resonance energy transfer (FRET) DNA melting assay, circular dichroism (CD) experiments and molecular dynamics simulations were used to gain insight into adduct formation and support HPLC/MS observations.

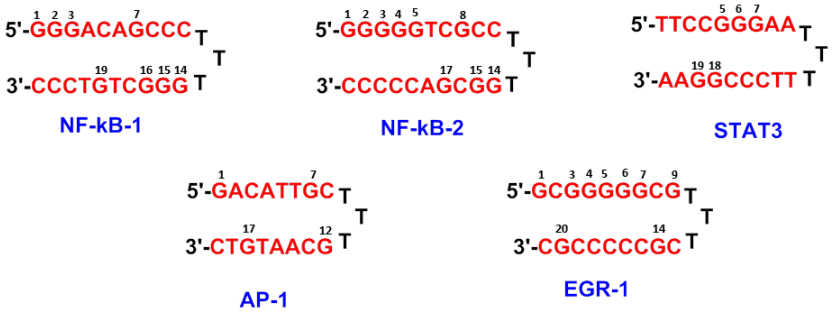


Figure 3.1: Hairpin oligonucleotides used in this study with the cognate sequences of the transcription factors NF-κB (two possible sequences were studied, in this study referred to as NF-κB-1 and NF-κB-2), EGR-1, AP-1 and STAT3 highlighted in red.

3.2 HPLC/MS study

Reversed phase HPLC has been used to study the extent and rate of adduct formation between SJG-136 and the DNA sequences depicted in **Figure 3.1**. According to the method developed in Chapter 2 (Method Ic), first each hairpin sequence was injected onto the HPLC to obtain reference peaks. SJG-136 was then incubated with each hairpin DNA and subjected to HPLC analysis. In the case of adduct formation new peaks emerge in the chromatographic profiles in addition to the parent DNA sequence which differ in their RTs compared to the DNA reference peaks (see error analysis in Section 2.4.1.2. In a similar fashion, each hairpin sequence was subjected to MALDI-TOF analysis to obtain reference peaks. The SJG-136/DNA mixture was subsequently subjected to MALDI-TOF-MS for stoichiometry identification. Similar to the MS spectra obtained in Chapter 2, the parent DNA appears with high intensity in the MS spectra of the incubation mixture. This is discussed in Section 2.4.1.2 in detail. Additional HPLC chromatograms

(5min incubation) and MALDI-TOF spectra (5 min and 3 hours incubation) are shown in the **Appendix**.

3.2.1 Results and discussion

NF-κB-1 sequence (**Figure 3.1**) was annealed according to the procedure described in Section 7.2.1.2. NF-κB-1 sequence alone gave a single peak at RT 7.069 min (**Figure 3.2A**) and a m/z of 7033.1 m/z (theoretical mass: 7032.6 m/z). After incubation with SJG-136 at 25 ºC in a 4:1 ratio (SJG-136/DNA) for various time intervals, two new minor peaks gradually emerged after 3 hours at RT 7.436 min and RT 8.169 min with approximately 16.6% adduct formed (**Figures 3.2B** and **C**). The 1:1 stoichiometry of the two new peaks was confirmed by MALDI-TOF-MS (observed mass: 7590.6 m/z; theoretical mass: 7589.2 m/z) as the SJG-136/NF-κB-1 adduct (**Figure 3.2D**). Based on the detected mass, the formed adduct was assumed to be the 1:1 SJG-136/NF-κB-1 adduct as the binding of the ligand to one guanine prevents further ligand association to the DNA sequence.

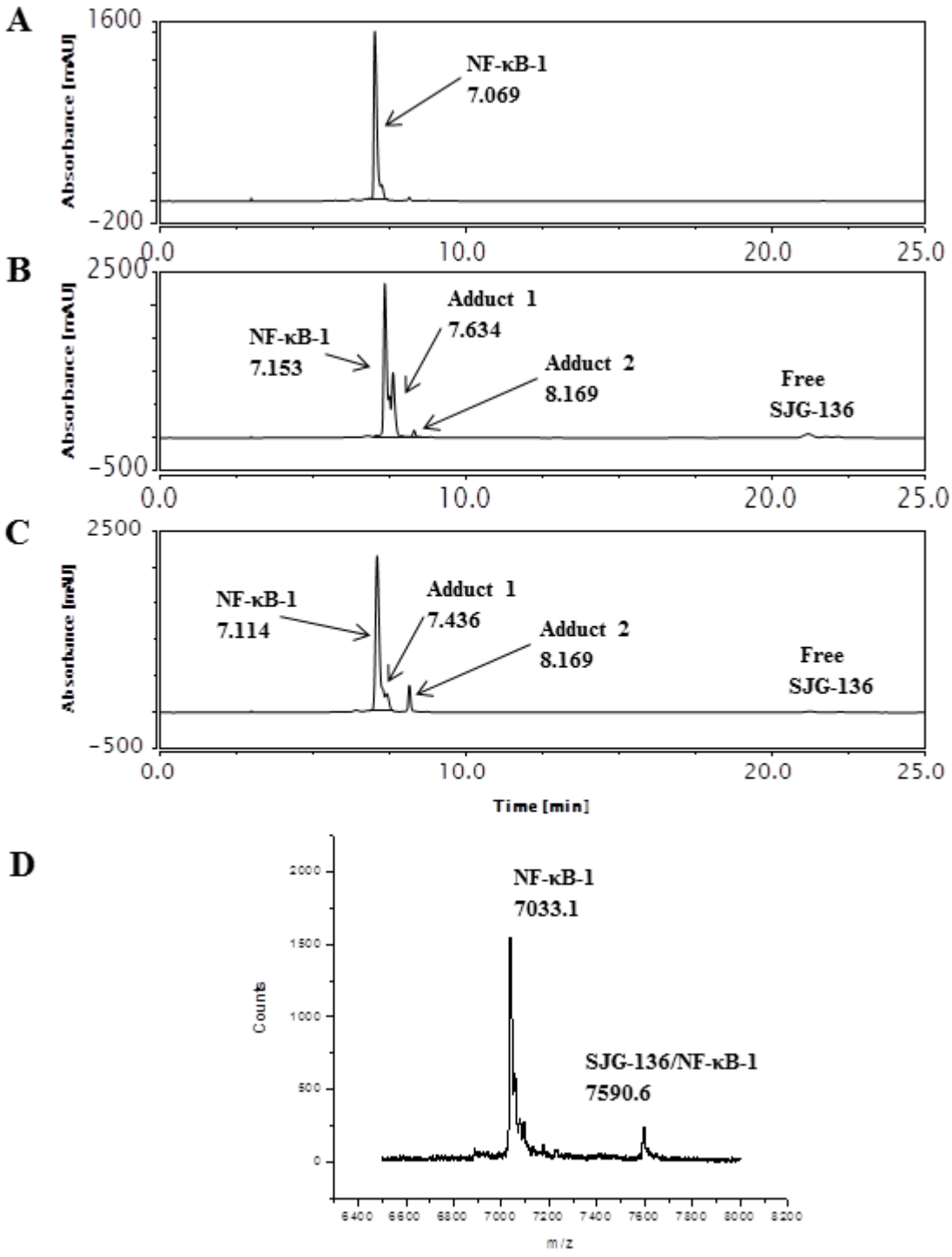


Figure 3.2: A, HPLC chromatogram showing the annealed NF-κB-1 sequence at RT 7.069 min; **B**, Annealed NF-κB-1 sequence after incubating with SJG-136 for 3 hours; **C**, annealed NF-κB-1 sequence after incubating with SJG-136 for 24 hours showing the appearance of two minor peaks at RT 7.436 min and RT 8.169 min with the reaction incomplete; **D**, MALDI-TOF spectrum of SJG-136/NF-κB-1 (after a 24 hours incubation) confirming the identity of adduct formation (NF-κB-1 observed mass: 7033.1 m/z, theoretical mass: 7032.6 m/z, NF-κB-1 adduct observed mass: 7590.6 m/z, theoretical mass: 7589.2 m/z).

Next, NF-κB-2 sequence (**Figure 3.1**) was annealed according to the procedure described in Section 7.2.1.2. Injection onto the HPLC column of NF-κB-2 sequence alone resulted in the appearance of a single peak at RT 7.097 min (**Figure 3.3A**) which was identified by MALDI-TOF-MS (7033.1 m/z). Following incubation with SJG-136 at 25 ºC in a 4:1 ratio (SJG-136/DNA) a new minor peak at RT 7.223 min emerged during the time-course study accounted for approximately 43% after the 24 hours incubation (**Figures 3.3B** and **C**). The 1:1 stoichiometry of the SJG-136/DNA adduct was confirmed by MALDI-TOF-MS (observed mass: 7590.7 m/z, theoretical mass: 7590.2 m/z) (**Figure 3.3D**). Similar to NF-κB-1 sequence, the observed mass of the formed adduct suggested that the 1:1 SJG-136/NF-κB-2 adduct was formed as the binding of the ligand prevents a further molecule from interaction with the NF-κB-2 sequence.

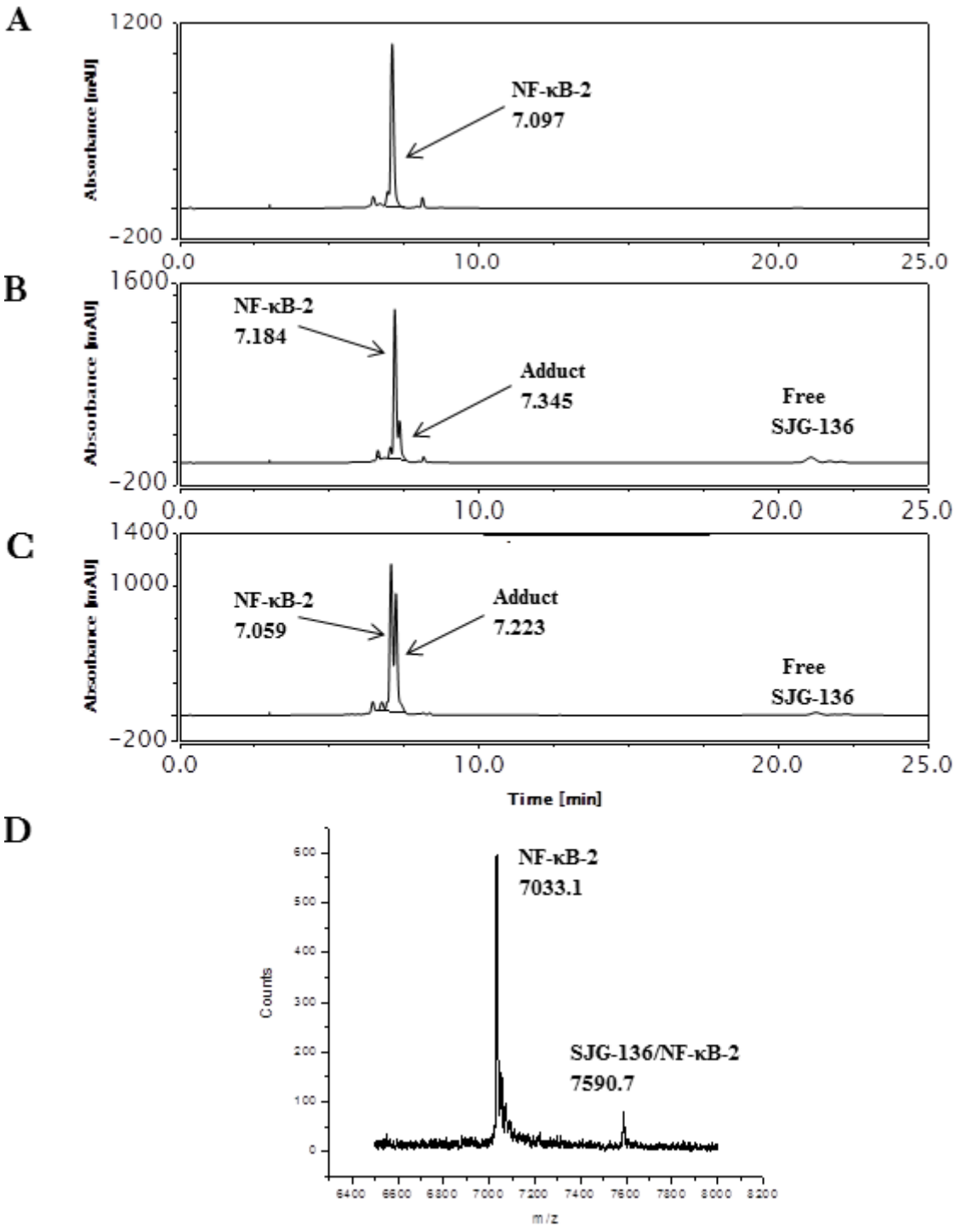


Figure 3.3: A, HPLC chromatogram showing the annealed NF-κB-2 sequence at RT 7.097 min; **B**, Annealed NF-κB-2 sequence after incubating with SJG-136 for 3 hours; **C** annealed NF-κB-2 sequence after incubating with SJG-136 for 24 hours showing the appearance of one minor adduct peak at RT 7.223 min with the reaction incomplete; **D**, MALDI-TOF spectrum of SJG-136/NF-κB-2 (after a 24 hours incubation) confirming the identity of adduct formation (NF-κB-2 observed mass: 7033.1 m/z, theoretical mass: 7033.6 m/z, NF-κB-2 adduct observed mass: 7590.7 m/z, theoretical mass: 7590.2 m/z).

EGR-1 sequence was annealed according to the procedure described in Section 7.2.1.2. In a similar experiment EGR-1 (**Figure 3.1**) was incubated with SJG-136 at 25 ºC in a 4:1 ratio (SJG-136/DNA) for different time points. Initially, EGR-1 sequence alone provided a single peak at RT 7.539 min (**Figure 3.4A**) with an observed mass of m/z 6416.9 (theoretical mass: 6416.2 m/z). After 3 hours incubation, a new minor peak at RT 8.469 min gradually emerged with approximately 28% adduct formation after 24 hours (**Figures 3.4B** and **C**). The 1:1 stoichiometry of the new peak was confirmed by MALDI-TOF-MS (observed mass: 6973.3 m/z, theoretical mass: 6972.8 m/z) as the SJG-136/EGR-1 adduct (**Figure 3.4D**). Analogous to NF-κB-1 and NF-κB-2 sequences, the detected mass of the formed adduct indicated that the ligand binds to the DNA and prevented another ligand molecule from further binding.

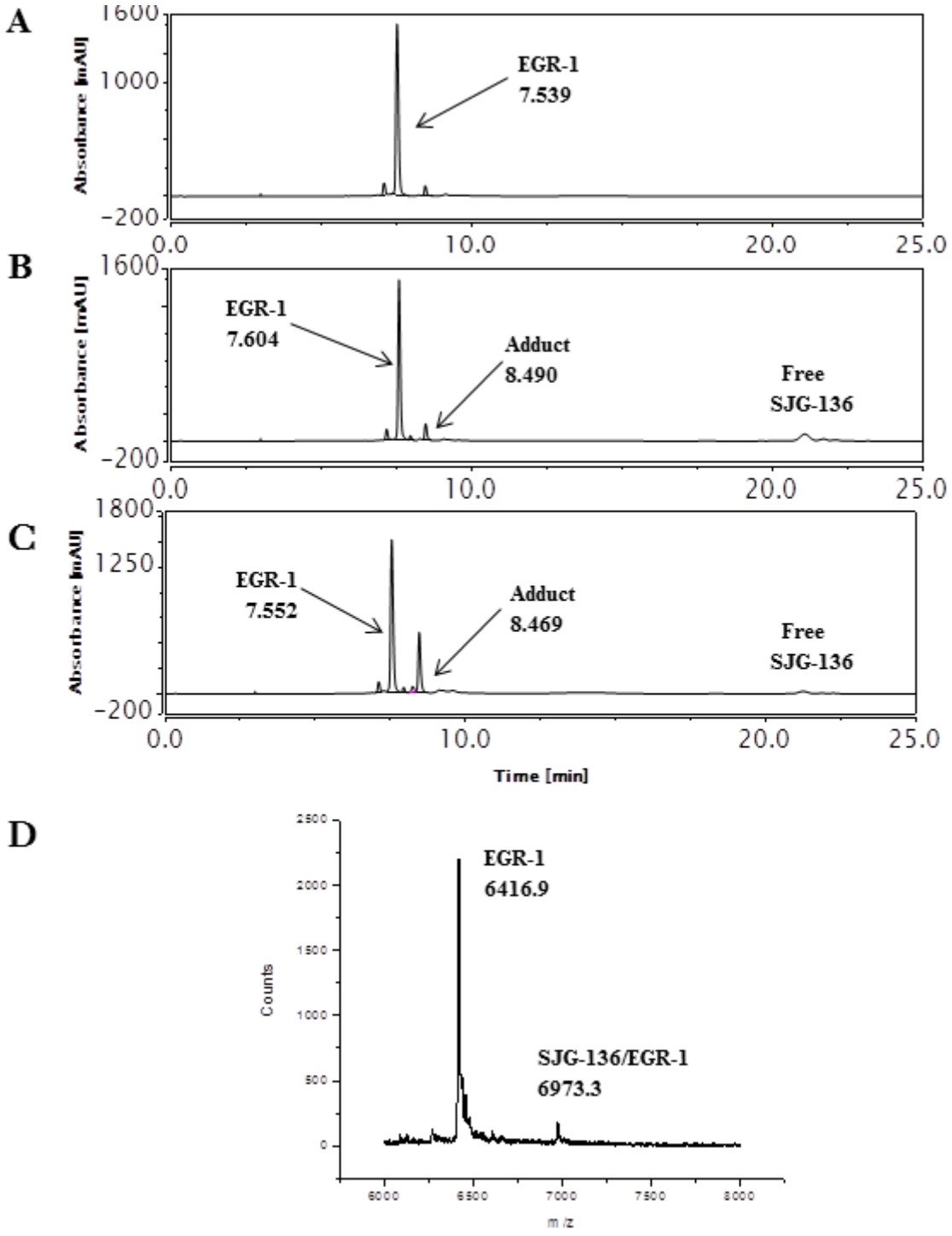


Figure 3.4: A, HPLC chromatogram showing the annealed EGR-1 sequence at RT 7.539 min; **B**, Annealed EGR-1 sequence after incubating with SJG-136 for 3 hours; **C** annealed EGR-1 sequence after incubating with SJG-136 for 24 hours showing the appearance of one minor adduct peak at RT 8.469 min with the reaction incomplete; **D**, MALDI-TOF spectrum of SJG-136/EGR-1 (after a 24 hours incubation) confirming the identity of adduct formation (EGR-1 observed mass: 6416.9 m/z, theoretical mass: 6416.2 m/z, EGR-1 adduct observed mass: 6973.3 m/z, theoretical mass: 6972.8 m/z).

AP-1 sequence was annealed according to the procedure described in Section 7.2.1.2. Annealed AP-1 sequence (**Figure 3.1**) gave a single peak in the HPLC chromatogram at RT 7.306 min (**Figure 3.5A**) identified by MALDI-TOF-MS. Incubation with SJG-136 at 25 °C in a 4:1 ratio (SJG-136/DNA) resulted in a rapid appearance of a new major peak at RT 7.494 min after 3 hours with 100% adduct formation (**Figure 3.5B**). No further changes were observed in the chromatographic profiles following the 3 hours incubation indicating reaction had gone to completion (**Figure 3.5C**). The stoichiometry of the adduct was confirmed as 1:1 SJG-136/AP-1 by MALDI-TOF-MS with an observed mass of 6351.2 m/z (theoretical mass: 6350.4 m/z) (**Figure 3.5D**). Similar to the other sequences, the detected mass of the formed adduct indicated 1:1 SJG-136/AP-1 adduct formation.

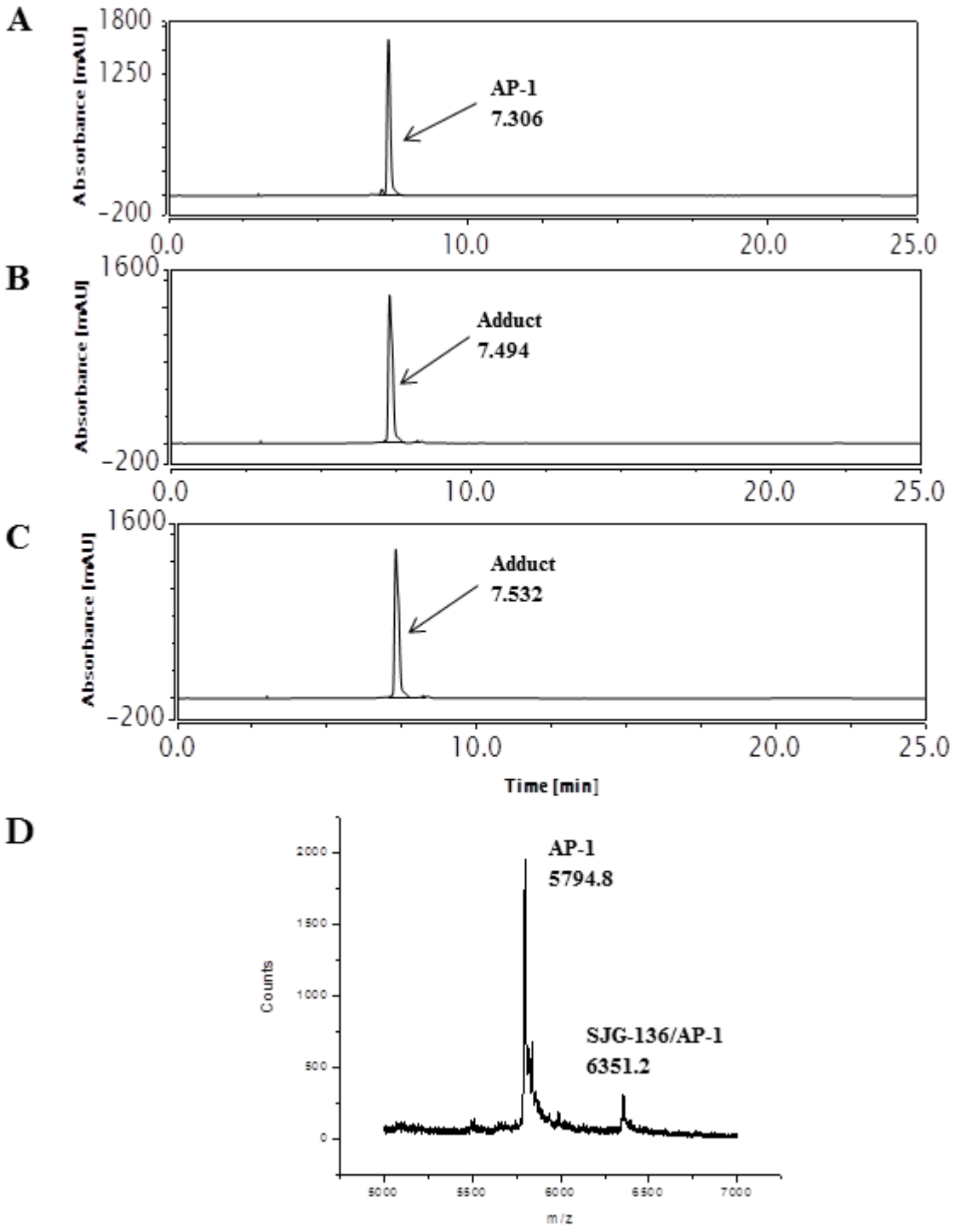


Figure 3.5: A, HPLC chromatogram showing the annealed AP-1 sequence at RT 7.306 min; **B**, Annealed AP-1 sequence after incubating with SJG-136 for 3 hours showing the appearance of a new major peak at RT 7.494 min with the reaction complete; **C** annealed AP-1 sequence after incubating with SJG-136 for 24 hours showing no further changes with the adduct peak at RT 7.532 min remaining; **D**, MALDI-TOF spectrum of SJG-136/AP-1 (after a 24 hours incubation) confirming the identity of adduct formation (AP-1 observed mass: 5794.8 m/z, theoretical mass: 5793.8 m/z, AP-1 adduct observed mass: 6351.2 m/z, theoretical mass: 6350.4 m/z).

STAT3 sequence was annealed according to the procedure outlined in Section 7.2.1.2. HPLC analysis of the STAT3 sequence (**Figure 3.1**) gave a single peak at RT 8.452 min (**Figure 3.6A**) which provided the correct m/z (6413.7) for this oligonucleotide by MALDI-TOF-MS. Following incubation with SJG-136 at 25 °C in a 4:1 ratio (SJG-136/DNA), three new minor peaks appeared in the HPLC chromatogram at RT 9.314 min, 10.296 min and RT 11.985 with approximately 14% adduct formation after 24 hours (**Figure 3.6B** and **C**). The adduct stoichiometry was confirmed by MALDI-TOF-MS as 1:1 SJG-136/STAT3 based on an observed mass of 6968.9 m/z (theoretical mass: 6968.8 m/z) (**Figure 3.6D**). The observed mass of a 1:1 SJG-136/STAT3 adduct indicated that the bound SJG-136 molecule prevents another from associating with the STAT3 sequence.

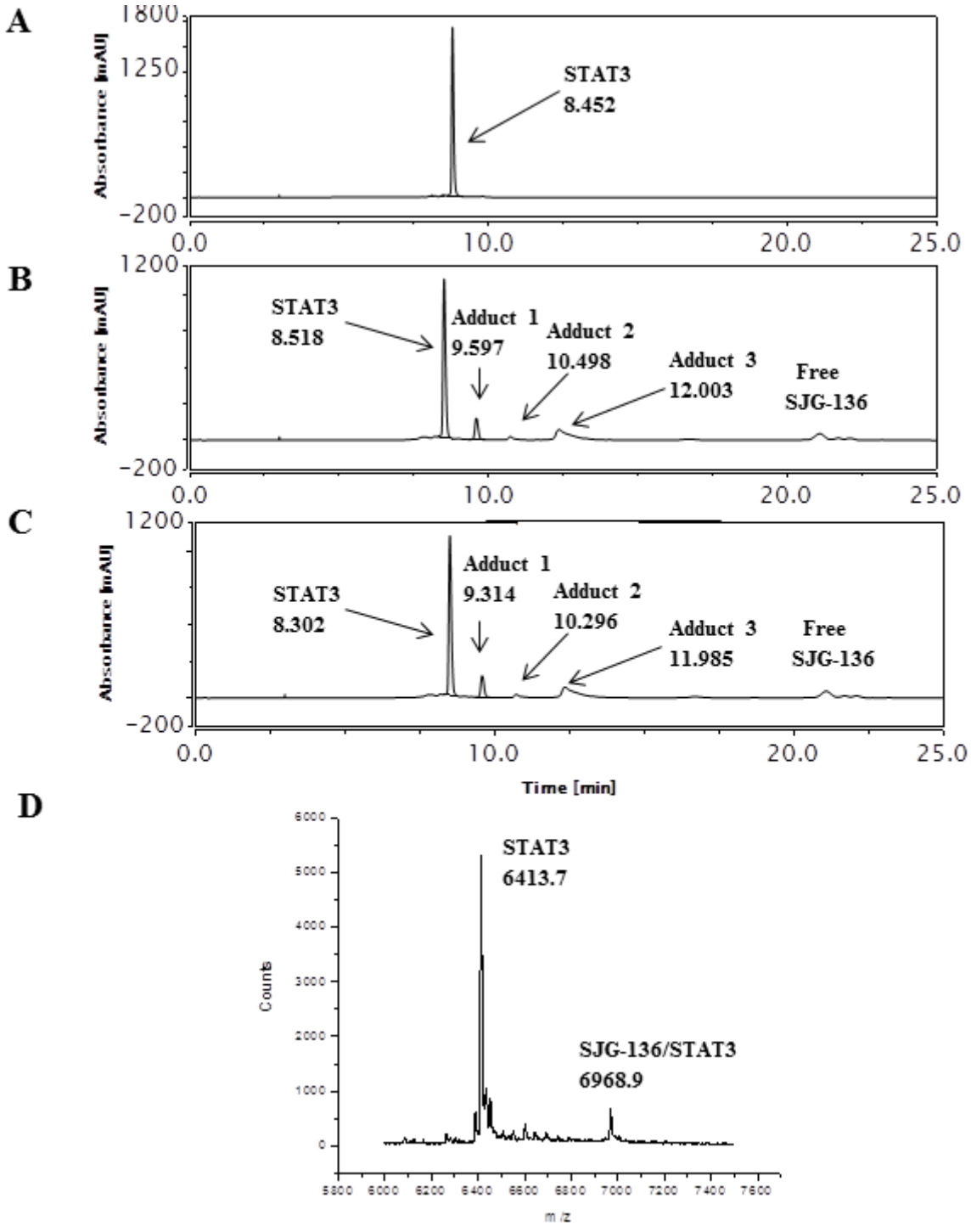


Figure 3.6: **A**, HPLC chromatogram showing the annealed STAT3 sequence at RT 8.452 min; **B**, Annealed STAT3 sequence after incubating with SJG-136 for 3 hours; **C** annealed STAT3 sequence after incubating with SJG-136 for 24 hours showing the appearance of three minor adduct peaks at RT 9.314 min, RT 10.296 min and RT 11.985 with the reaction incomplete; **D**, MALDI-TOF spectrum of SJG-136/STAT3 (after a 24 hours incubation) confirming the identity of adduct formation (STAT3 observed mass: 6413.7 m/z, theoretical mass: 6412.2 m/z, STAT3 adduct observed mass: 6968.9 m/z, theoretical mass: 6968.8 m/z).

The results obtained from the HPLC/MS study indicate that SJG-136 may be capable of binding to NF-κB, EGR-1, AP-1 and STAT3 consensus sequences. Based on the developed HPLC method, adduct formation was observed for all studied sequences, providing the correct molecular weight for the 1:1 SJG-136/DNA adducts by MALDI-TOF-MS (**Table 3.1**).

Table 3.1: Theoretical and observed masses [Da] of single-stranded hairpin oligonucleotides and their 1:1 adducts with SJG-136 after a 24 hours incubation.

DNA Sequence	DNA Mass	Theoretical Mass DNA/SJG-136 (1:1) (DNA Mass + 556.64)	Observed DNA/SJG-136 Adduct Mass (1:1)
NF-κB-1	7032.6	7589.2	7590.6
NF-κB-2	7033.6	7590.2	7590.7
EGR-1	6416.2	6972.8	6973.3
AP-1	5793.8	6350.4	6351.2
STAT3	6412.2	6968.8	6968.9

Surprisingly, the data revealed significant differences in extent of adduct formation between the individual sequences although their guanine content is comparable (**Table 3.2**). SJG-136 rapidly formed an adduct with NF-κB-2 sequence (~43% adduct formed after 24 hours) and AP-1 sequence (~100% adduct formation after 5 min) while it reacted moderately with the EGR-1 sequence (~28% adduct formed after 24 hours), slowly with the NF-κB-1 sequence (~16% adduct formed after 24 hours) and STAT3 sequence (~14% adduct formation after 24 hours) (**Figure 3.7**In a similar fashion, the extent of adduct formation starkly contrasted between the NF-κB-1 and NF-κB-2 sequences although both sequences differ in only one guanine, eight and nine respectively. Interestingly, SJG-136 appeared to be highly reactive towards the AP-1 sequence, containing four guanines while it has shown low reactivity towards the STAT3 sequence which comprises five guanines. However, similar to the observation made on the interaction between GWL-78 and KMR-28-39 (see Sections 2.4.1.2 and 2.4.1.3) the difference in the RT for the AP-1 sequence

and the corresponding adduct was observed to be very small. This minimal difference makes it difficult to state that SJG-136 reacted with sequence AP-1 to full extent after 5 min. This observation may be due to the insufficient resolution of the solvent system used in this method. Therefore, SJG-136 may have not reacted to the full extent with the AP-1 sequence, but due to the similar polarities of the parent AP-1 sequence and SJG-136/AP-1 adduct they elute simultaneously, resulting in one major peak. Furthermore, as for GWL-78 and KMR-28-39, very small differences in the RTs between DNA (*e.g.*, NF-κB-1, NF-κB-2 and AP-1) and SJG-136/DNA adducts were observed. This has been discussed in Section 2.4.1.2 in detail. As in the case of GWL-78 and KMR-28-39, all MALDI-TOF spectra showed the appearance of the parent DNA. This has been discussed in Section 2.4.2.1 in detail.

Table 3.2: Extent of adduct formation after 24 hours SJG-136 with transcription factor sequences NF-κB-1, NF-κB-2, EGR-1, AP-1 and STAT3 (formed adduct in %).

	NF-κB-1	**NF-κB-2**	**EGR-1**	**AP-1**	**STAT3**
SJG-136	16.61	42.90	27.73	100	13.94

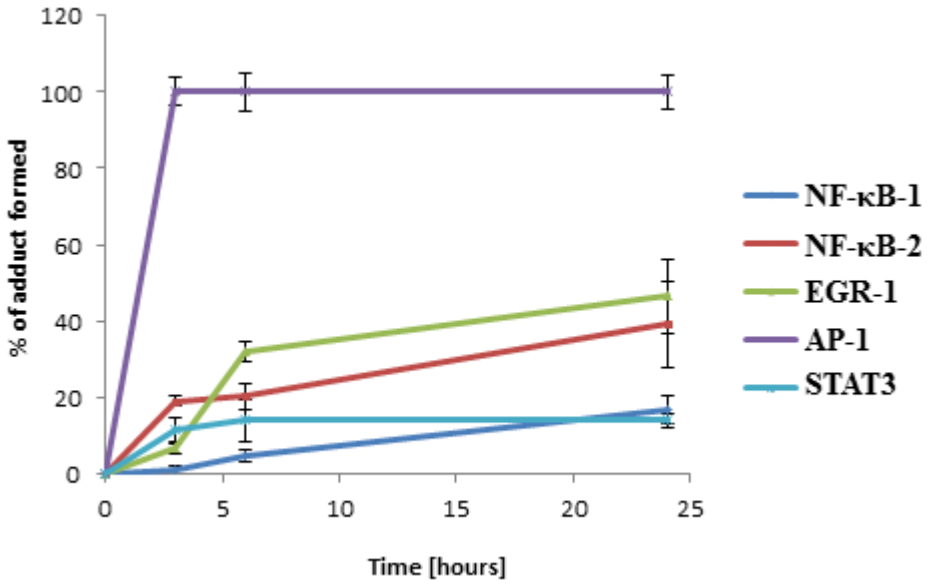


Figure 3.7: Graphical representation showing variability in adduct formation between SJG-136 and different transcription factor binding sequences NF-κB-1, NF-κB-2, EGR-1, AP-1 and STAT3.

Equally important are the observed differences in the rate of adduct formation (**Table 3.3**). Previous studies have demonstrated that SJG-136 can form interstrand crosslinks at Pu-GATC-Py, longer interstrand cross-links at Pu-GAATC-Py, intrastrand cross-linked adducts at Pu-GATG-Py, and longer intrastrand links at Pu-GAATG-Py sequences (rank order: Pu-GAAT > Pu-GATC-Py >> Pu-GATG-Py > Pu-GAATC). In addition, SJG-136 is capable of forming mono-alkylated adducts with only one PBD unit covalently bound at sequences that contain suitable PBD binding sites but where neither inter- nor intra-strand cross-links are feasible due to the unavailability of two appropriately positioned guanines[96]. Surprisingly, on the basis of our HPLC results SJG-136 formed three distinct adducts with the STAT3 sequence, two different adduct types with the NF-κB-1 sequence, while only single adducts were determined. In the case of NF-κB-2, EGR-1 and AP-1 sequences. These results suggest that SJG-136 favours one particular binding site within NF-κB-2, EGR-1 and AP-1 sequences as only one particular type of adduct had been formed and further strengthen the sequence selectivity of SJG-136.

Table 3.3: Number of adducts observed after 24 hours between SJG-136 and the transcription factor sequences NF-κB-1, NF-κB-2, EGR-1, AP-1 and STAT3 (number of adducts formed).

	NF-κB-1	**NF-κB-2**	**EGR-1**	**AP-1**	**STAT3**
SJG-136	2	1	1	1	3

With the NF-κB-1 sequence, SJG-136 can form various types of adducts such as mono-alkylated adducts, inter- and intrastand cross-links as this sequence contains multiple binding sites for SJG-136. Within NF-κB-1 sequence mono-adducts can be formed at G2, G3, G7, G15, G16 and G19. An interstrand cross-link is likely to appear between G7 and G19. Lastly, G3 and G7 are favourable for an intrastrand interaction between SJG-136 and NF-κB-1 sequence. According to previous reports, guanines located at the 5'- or 3'-end are not suitable binding sites for SJG-136 as the ligand would have to protrude beyond the 5'- or 3'-end of the hairpin with a loss of DNA/ligand stabilising interactions[96], and therefore G1 is unfavourable for SJG-136 binding. Furthermore, the interaction between SJG-136 and the guanine at G14 are not expected to occur due to proximity of the guanine base to the TTT-loop which will result in steric interference. Although within the NF-κB-1 sequence various guanines are available for a covalent attachment of SJG-136, we have observed that only two particular types of adducts had been formed suggesting two preferred binding sites of SJG-136.

SJG-136 can bind to sequence NF-κB-2 and form mono-alkylated adducts at G2, G3, G4, G5, G8, G15 and G17. An interstrand cross-link is likely to appear between G5 and G15. G4, G17 and G5 and G8 are suitable binding sites for an intrastrand interaction between SJG-136 and NF-κB-2 sequence. Similarly, to NF-κB-1 sequence, within NF-κB-2 sequence G1 is not a favourable binding position for SJG-136 due to its 5'-end location as well as G14 which resides next to the TTT-loop. It was surprising given that, although many different adduct types may be formed between SJG-136 and NF-κB-2 sequence only one particular adduct had been formed. Moreover, the data suggests that SJG-136 has a higher affinity toward the NF-κB-2 sequence compared to NF-κB-1 sequence although their sequence and guanine content is similar.

Analogous to NF-κB-1 and NF-κB-2 sequences, EGR-1 sequence is also guanine rich and contains several suitable binding sites for SJG-136. Mono-alkylation within EGR-1 sequence may occur at G3, G4, G5, G6, G7, G14 and G20. An interstrand cross-linked adduct may be formed between G14 and G4 or G5, and G20 and G6 or G7. Finally, G3 and G7 are favourable positions for an intrastrand cross-linked adduct. Presumably, SJG-136 does not to bind to G1 located at the 5'-end of EGR-1 sequence as this would require the ligand to protrude beyond the sequence in a similar fashion to NF-κB-1 and NF-κB-2 sequences. Furthermore, binding to G9 is not expected as steric interference may appear due to the proximate location of the guanine to the TTT-loop. In accordance with observations made for NF-κB-1 and NF-κB-2 sequences, SJG-136 formed only a single type of adduct with the EGR-1 sequence despite various favourable binding sites. This is again further support for the sequence selectivity and specific binding profile of SJG-136.

The most striking results were obtained for AP-1 with 100% adduct formation after 5 min incubation with SJG-136 and one main adduct formed (**Figure 3.5** and **3.7**). This suggests that SJG-136 favours a specific binding site within AP-1 sequence due to the observed extent and rate of adduct formation compared to the NF-κB, EGR-1 and STAT3 sequences (**Figure 3.7**). Similar to the other sequences, also with the AP-1 sequence SJG-136 can form various adducts. Mono-alkylation within the AP-1 sequence may occur at G7 and G17. Interstrand cross-links can be formed at G7 and G17. It is not expected (based on previous literature results) that SJG-136 interacts with G1 at the 5'-end of AP-1 sequence for the same reasons discussed for NF-κB and EGR-1 sequences. Additionally, G12 is not a favourable binding site for the ligand as steric interference may occur due to its nearby location to the TTT-loop.

The rapid rate of adduct formation within AP-1 sequence led us to investigate in detail the binding site of SJG-136 in order to identify the guanine which is responsible for the binding behaviour of SJG-136. To address this issue, the guanines were subsequently replaced with inosine bases which lack the nucleophilic NH_2 group and, therefore, prevent SJG-136 from binding to the sequence. The results obtained from this study are presented in Chapter 5.

SJG-136 demonstrated low reactivity toward STAT3 sequence due to slow reaction (approximately 14% adduct formed after 24 hours incubation) and the appearance of

multiple adducts (**Figure 3.6** and **3.7**). Mono-alkylation within STAT3 sequence may occur at G5, G6, G7, G18 and G19ss. Interstand cross-linked adducts may be formed between G18 and G6 or G7 and G7 and G18.

In order to rationalise the type of adducts formed with sequences NF-κB, EGR-1, AP-1 and STAT3, molecular dynamics simulations and free energy binding calculations were undertaken, and these will be discussed in greater detail in Section 3.5.

3.2.2 Summary of HPLC/MS study

Taken together, the HPLC/MS study has indicated that SJG-136 may be capable of binding to the cognate sequences of the transcription factors NF-κB, EGR-1, AP-1 and STAT3. Adduct formation was observed during the HPLC study for all studied sequences. The stoichiometry of the adducts formed was subsequently confirmed by MALDI-TOF-MS. Interestingly, the data revealed significant differences in rate and extent of adduct formation **Figure 3.7**, **Table 3.2** and **3.3**. The ligand appeared to react rapidly with the AP-1 (~100% adduct formation after 24 hours), moderately with the NF-κB-2 (~43% adduct formed after 24 hours) and EGR-1 (~28% adduct formation after 24 hours) and slowly with the NF-κB-1 (~17% adduct formed after 24 hours) and STAT3 (~14% adduct formation after 24 hours) (**Figure 3.7**). Limitations of the developed method and their effects on the results obtained for the AP-1 sequence have been discussed above. Moreover, multiple adducts have been detected for STAT3 and NF-κB-1 sequences, three and two, respectively whereas only one type of adduct was formed with the NF-κB-2, EGR-1 and AP-1 sequences (**Table 3.3**) although multiple suitable binding sites for SJG-136 are available within these sequences. This result suggests a preferred binding site of SJG-136 within NF-κB-2, EGR-1 and AP-1 sequences.

The observed differences in reactivity of SJG-136 toward the consensus sequences of the transcription factors NF-κB, EGR-1, AP-1 and STAT3 had previously not been reported and were assumed not to vary in this extent. The high reactivity of SJG-136 toward AP-1 sequence was particularly surprising, as it means that the ligand must favour AP-1 cognate sequence over NF-κB, EGR-1 and STAT3. From these results it was concluded that SJG-136 exhibits better sequence-selectivity than originally thought as it appeared to react to full extent with one adduct formed with the AP-1 consensus sequence.

Taken together, the HPLC/MS study has indicated that SJG-136 may be able to bind to the consensus sequences of NF-κB, EGR-1, AP-1 and STAT3 transcription factors. Remarkable differences in extent and rate of adduct formation have been observed for individual sequences. This surprising result is of utmost importance as it adds new data on the overall mechanism of action of SJG-136, and in addition can help to understand the previously reported differences in activity of SJG-136 in various human tumour cell lines.

3.3 FRET study

The DNA-binding affinity of SJG-136 toward the NF-κB-1, NF-κB-2, EGR-1, AP-1 and STAT3 sequences was examined using the FRET DNA melting assay. DNA binding agents stabilise the secondary structure of DNA upon binding which results in an increase in the melting temperature of the ligand-DNA complex. The differences in the melting temperature of the DNA/ligand complex are compared with the melting temperature of the DNA alone and expressed as ΔT_m. During the FRET experiment the process of melting is monitored by measuring the fluorescence of the FRET labelled DNA (**Figure 3.8**).

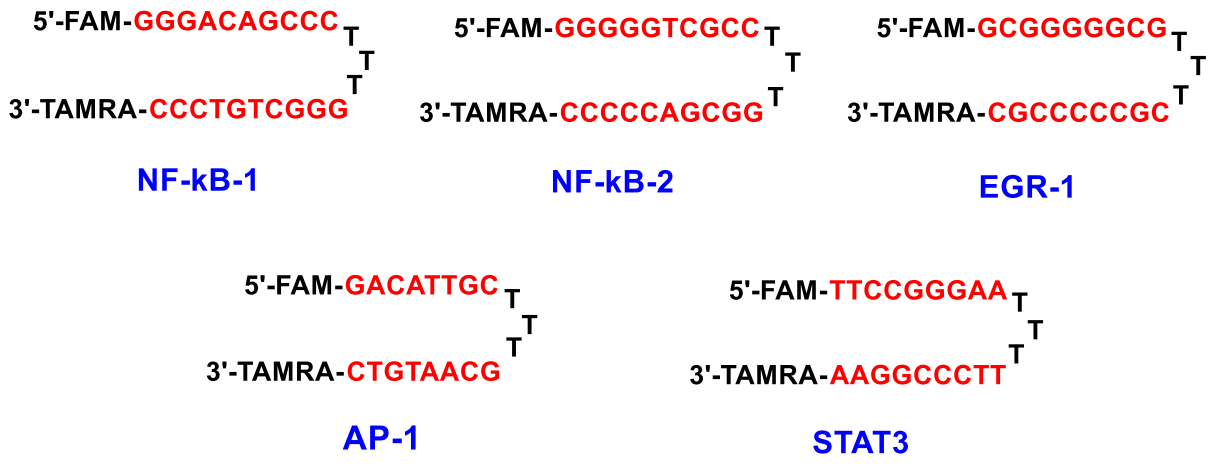


Figure 3.8: Structure of FRET labelled hairpin oligonucleotides used in the DNA melting assay.

3.3.1 Results and discussion

Working solutions of DNA, SJG-136 and SJG-136/DNA complexes were prepared according to the procedure described in Section 7.2.5. The melting temperatures of each sequence before treatment with SJG-136 are shown in **Table 3.4**. All FRET melting curves are shown in the **Appendix**. The melting temperatures of the NF-κB-1, NF-κB-2 and EGR-1 sequences are very high (>75 °C). Therefore, it is difficult to interpret the FRET data for the NF-κB-1, NF-κB-2 and EGR-1 sequences. However, the melting temperatures of the AP-1 and STAT3 show appropriate melting values.

Table 3.4: Average melting temperature [°C] of the hairpin sequences alone during the FRET study.

NF-κB-1	NF-κB-2	EGR-1	AP-1	STAT3
78.5	83.4	83.1	67.6	60.8

For NF-κB-1 sequence, SJG-136 exhibited greater binding affinity at higher concentrations of 50 μM and 100 μM with an observed ΔT_m of 17.6 °C and was less effective in stabilising NF-κB-1 sequence at lower concentrations of 0.1 μM, 1.0 μM and 10 μM with an observed increase in ΔT_m of < 1 °C) (**Table 3.5**). The similar ΔT_m values at 50 μM and 100 μM suggest saturation of binding to NF-κB-1 at 50 μM.

Table 3.5: ΔT_m [°C] differences of NF-κB-1, NF-κB-2, EGR-1, AP-1 and STAT3 sequences after incubation with SJG-136 for 24 hours at various concentrations. Data are means of two readings.

SJG-136 [μM]	NF-κB-1	NF-κB-2	EGR-1	AP-1	STAT3
100	17.62	11.6	10.63	20.22	4.57
50	17.6	11.32	6.66	20.54	3.17
10	0.83	2.45	1.14	20.5	1.62
1	0.95	0.95	0.25	0.865	0.77
0.1	0.465	1.1	1	0.6	0.84

Similar stabilising effect of SJG-136 was observed for the NF-κB-2 sequence (**Table 3.5**). At ligand concentrations of 50 µM and 100 µM the ΔT_m increased up to ~11 ºC with no stabilisation effect upon ligand binding at lower concentrations (**Figure 3.9**). Compared to NF-κB-1 sequence, SJG-136 demonstrated a lower stabilising effect on NF-κB-2 sequence at higher concentrations of 50 µM and 100 µM (NF-κB-1 sequence ΔT_m = 17 ºC, NF-κB-2 sequence ΔT_m = 11 ºC). Moderate stabilisation of NF-κB-2 sequence was achieved at 10 µM SJG-136 with ΔT_m of 3.27 ºC compared to NF-κB-1 sequence with ΔT_m of 0.8 °C at the same ligand concentration.

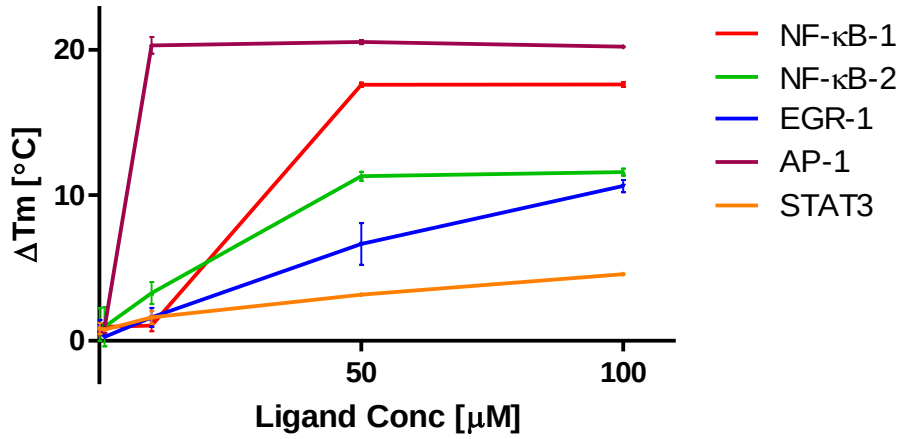


Figure 3.9: Graphical representation of ΔT_m [°C] differences of NF-κB-1, NF-κB-2, EGR-1, AP-1 and STAT3 sequences after incubation with SJG-136 for 24 hours at various concentrations.

Analogous to NF-κB-1 and NF-κB-2 sequences an increase in ΔT_m of the formed SJG-136/EGR-1 complex was observed at SJG-136 concentrations of 50 µM (ΔT_m ~6.66 ºC) and 100 µM (ΔT_m ~10.63 ºC) (**Table 3.5**). In contrast to NF-κB-1 and NF-κB-2 sequences no saturation of DNA binding appeared at 50 µM as the ΔT_m of ~6.66 ºC at 50 µM elevated to ~10.63 ºC at 100 µM. Similar to NF-κB-1 sequence, at low concentrations of 0.1 µM, 1.0 µM and 10 µM SJG-136 exhibited low DNA binding ability with no increase in ΔT_m (**Table 3.5** and **Figure 3.9**).

The FRET results have demonstrated that SJG-136 exhibited the greatest stabilisation effect on AP-1 sequence (**Table 3.5** and **Figure 3.9**). An increase in the melting temperature of the SJG-136/AP-1 complex of ~20 ºC was observed already at SJG-136

concentrations of 10 µM. This contrasted starkly with the ΔT_m values obtained for NF-κB-1, NF-κB-2, EGR-1 and STAT3 sequences where a ligand induced DNA stabilisation effect was achieved starting from 50 µM. The ΔT_m values for AP-1 sequence at 10 µM, 50 µM and 100 µM SJG-136 remained at a similar value of ~20 ºC indicating maximum binding of SJG-136 to AP-1 sequence at these concentrations. Similar to NF-κB-1, NF-κB-2, EGR-1 and STAT3 sequences SJG-136 exhibited a very little stabilising effect on AP-1 sequence at low concentrations of 0.1 µM and 1.0 µM. The FRET results were in broad agreement with our observations made during the HPLC/MS study where SJG-136 rapidly formed one main adduct with the AP-1 hairpin sequence. These data further support the observed preference of SJG-136 for the AP-1 consensus sequence.

During the FRET study SJG-136 was demonstrated to have the lowest stabilising ability on STAT3 sequence compared to NF-κB-1, NF-κB-2, EGR-1 and AP-1 sequences (**Table 3.5** and **Figure 3.9**). Moderate elevation of the SJG-136/STAT3 complex melting temperature was measured at 50 µM and 100 µM SJG-136 with ΔT_m ~3.17 ºC and ~4.57 ºC, respectively whereas for the other sequences the ΔT_m values at 100 µM SJG-136 increased at least up to 10 ºC (**Table 3.5**). As anticipated, no remarkable increase in ΔT_m could be determined at 0.1 µM, 1.0 µM, and 10 µM SJG-136. These results correlated well with the HPLC/MS observations where the ligand was found to be less reactive and selective toward the STAT3 sequence.

3.3.2 Summary of the FRET study

In summary, from the FRET data listed above, SJG-136 appeared to stabilise NF-κB-1, NF-κB-2, EGR-1 and STAT3 sequences to some extent at higher concentrations of 50 µM and 100 µM (**Table 3.5** and **Figure 3.9**). No DNA stabilisation was achieved at lower ligand concentrations of 10 µM, 1 µM and 0.1 µM (**Table 3.5** and **Figure 3.9**). Several questions remained however. It is difficult to analyse the FRET data as some hairpins (*e.g.*, NF-κB-1, NF-κB-2 and EGR-1) melted at such high temperatures (>75 °C) that proper evaluation of DNA stabilisation couldn't be achieved. Furthermore, at around 75 °C, the degradation of some PBD-Guanine covalent bonds can occur which may result in false negative melting data. Therefore, the data obtained from the FRET study should be taken with caution as this experiment was used as a preliminary screen. Interestingly, SJG-136 has demonstrated a significant elevation in the melting temperature of AP-1

sequence already at 10 μM (ΔT_m of ~ 20 °C) (**Table 3.5** and **Figure 3.9**). This result suggests a greater stabilisation effect on the AP-1 sequence due to ligand binding compared to the other sequences that may be explained through the presence of a favourable binding site for SJG-136 within AP-1 sequence. The FRET results for the sequence AP-1 were in broad agreement with the HPLC/MS experiments, in which selectivity for the AP-1 sequence has been observed based on the fastest reaction with only one main adduct formed. Additionally, they further support the assumption that SJG-136 may exhibit its anti-tumour activity also through TF inhibition. As discussed above, the data obtained for the NF-κB-1, NF-κB-2 and EGR-1 sequences should be analysed carefully for the reasons mentioned above.

3.4 CD study

CD spectroscopy was used to explore changes in the secondary structure of the DNA sequences upon SJG-136 binding. Working solutions of DNA (final concentration of 5 μM), SJG-136 (final concentration of 20 μM) and SJG-136/DNA complexes were prepared according to the protocol described in Section 7.2.5. Initially, CD spectra of the DNA sequences (**Figure 3.1**) alone were recorded between 200 and 400 nm in order to obtain reference spectra for subsequent comparison after SJG-136 addition. Following this, the DNA sequences were incubated with SJG-136 in a 4:1 ratio (SJG-136/DNA) and either subjected immediately to CD measurements for t = 0 h experiments or incubated for 24 hours at 25 ºC followed by CD analysis as described in Section 7.2.6.

3.4.1 Results and discussion

The CD profiles of all sequences display a negative band at approximately 240 nm and a positive band at approximately 272 nm (**Figure 3.10**) and are typical CD profiles of the B-form hairpin DNA. It has been previously reported that the CD spectrum of pure B-form DNA exhibits a negative band at about 245 nm and positive bands at about 260-280 nm[178c, 204]. The negative band at about 245 nm appears due to the right-handed helicity of the DNA whereas the positive band at around 260-280 nm occurs due to the π-π base stacking. It should be stressed that the position and amplitudes of the CD bands may vary

with the sequence. This is due to differences in chromophores and also because of various conformational properties.

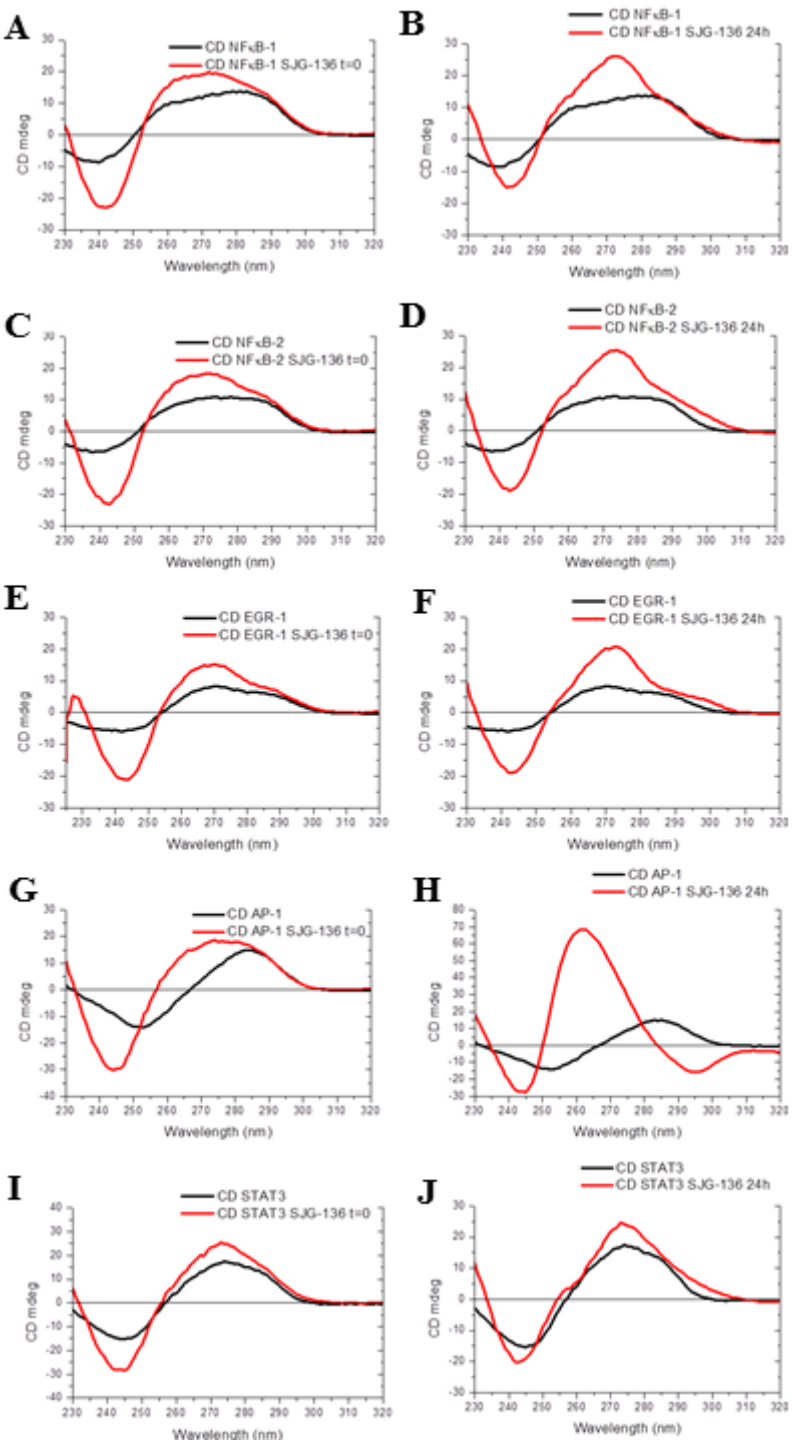


Figure 3.10: CD spectra: **A**, NF-κB-1 sequence alone (black) and NF-κB-1/SJG-136 complex at t = 0 h; **B**, NF-κB-1 sequence alone (black) and NF-κB-1/SJG-136 complex at t = 24 h showing binding of SJG-136 to NF-κB-1sequence**; C**, NF-κB-2 sequence alone (black) and NF-κB-2/SJG-136 complex at t = 0 h; **D**, NF-κB-2 sequence alone (black) and NF-κB-2/SJG-136 complex at t = 24 h confirming binding of SJG-136 to NF-κB-2 sequence; **E**, EGR-1 sequence alone (black) and EGR-1/SJG-136 complex at t = 0 h; **F**, EGR-1 sequence alone (black) and EGR-1/SJG-136 complex at t = 24 h showing binding of SJG-136 to EGR-1 sequence; **G**, AP-1 sequence alone (black) and AP-1/SJG-136 complex at t = 0 h; **H**, AP-1 sequence alone (black) and AP-1/SJG-136 complex at t = 24 h confirming binding of SJG-136 to AP-1 sequence; **I**, STAT3 sequence alone (black) and STAT3/SJG-136 complex at t = 0 h; **J**, STAT3 sequence alone (black) and STAT3/SJG-136 complex at t = 24 h showing binding of SJG-136 to STAT3 sequence.

The CD profiles of all sequences showed an increase in the negative ellipticity at about 240 nm and positive ellipticity at about 272 nm after addition of SJG-136 and immediate CD analysis (**Figure 3.10A**, **C**, **E**, **G** and **I**). These changes in the CD bands confirm alterations in the secondary structure of DNA upon SJG-136 binding.

Similar results were obtained after 24 hours incubation. The CD profiles of all sequences displayed an enhancement of the negative band at about 240 nm and the positive band at about 270 nm (**Figure 3.10B**, **D**, **F**, **H** and **J**). Moreover, the positive CD band increased to a greater extent for all sequences after 24 hours incubation with SJG-136 compared to the t = 0 experiment. This result reflects the time-dependent extent of adduct formation showing that more adduct had been formed after 24 hours.

Interestingly, after addition to AP-1 sequence, SJG-136 induced a significant increase of the positive CD band at about 263 nm from 4 mdeg up to 55 mdeg after 24 hours compared to t = 0 (**Figure 3.10G** and **H**). Furthermore, the AP-1/SJG-136 complex exhibited the biggest change in CD intensity of the positive band compared NF-κB-1, NF-κB-2, EGR-1 and STAT3 sequences indicating greater reactivity of SJG-136 toward AP-1 sequence. This result was consistent with our previous HPLC/MS and FRET data where the ligand has demonstrated a high preference and selectivity for AP-1 sequence.

3.4.2 Summary of CD study

Taken together, the CD study has confirmed the ability of SJG-136 to bind to the DNA target sequences of the TFs NF-κB, EGR-1, AP-1 and STAT3. All studied sequences displayed changes in their characteristic CD signals, demonstrating alterations in the secondary structure after SJG-136 binding which were particularly more pronounced after the 24 hours incubation. However, the results obtained from the t = 0 CD experiment carried out on AP-1 sequence does not support our HPLC/MS observation. During the HPLC/MS study, SJG-136 immediately formed an adduct with the AP-1 sequence (~100% adduct formation) after 5 minutes based on the presence of one main peak in the chromatographic profile. The HPLC chromatograms obtained during the time-course study did not change over time suggesting complete reaction after 5 minutes. However, from the observed differences in the CD spectra at t = 0 and 24 hours it becomes apparent that more adduct was formed after 24 hours indicating that the adduct formation was not complete after 5 minutes. These discrepancies in the HPLC and CD results after 5 min incubation are presumably due to the fact that the HPLC peak was not resolved and the parent AP-1 sequence and the formed adduct eluted from the column at the same time. Nevertheless, the CD study provided further support for the observations made on AP-1 sequence during the HPLC/MS and FRET study. In accordance with HPLC results, the AP-1/SJG-136 complex displayed the biggest change in the positive CD band especially after the 24 hours incubation intensity, compared to NF-κB-1, NF-κB-2, EGR-1 and STAT3 sequences, again demonstrating greater reactivity due to selectivity toward AP-1 sequence.

3.5 Molecular modeling study*

Molecular dynamics simulations were used to study the reactivity of SJG-136 toward NF-κB-1, NF-κB-2, EGR-1, AP-1 and STAT3 sequences in an effort to rationalise the results of HPLC/MS, FRET and CD studies. As SJG-136 is known to form mono-adducts and cross-links, both types of adduct were investigated, and in the case of mono-adducts, both loop-facing and non-loop-facing orientations were considered. Potential energy calculations (kcal/mol), averaged over the duration of a 10 ns simulation, were used to evaluate the binding potential of SJG-136 to each guanine in the DNA sequences. Based on publications to date three base-pairs were set as minimum criteria for adduct formation (based on the preference of the PBD for 5'-X-G-X-3' triplets). However, we have recently observed that SJG-136 is capable of binding to sequences while spanning only two base pairs. These results will be discussed in greater detail in Chapter 5.

**Molecular modeling studies were undertaken by Dr Paul Jackson*

3.5.1 Results and discussion

Potential energy calculations (kcal/mol) suggest that in mono-alkylated form, SJG-136 should preferentially bind to G3 of the NF-κB-1 consensus sequence (**Table 3.6**). The lowest energy snapshot of a 10 ns molecular dynamics simulation illustrates SJG-136 embedded in the minor groove of DNA, potentiated by non-covalent interactions between the central methylene linker of the dimer and the A4:T20 base-pair. Furthermore, as the second PBD of the dimer orientates over the A6:T18 base-pair, non-covalent interactions between the unreacted PBD and A:T assist in restraining the ligand in the minor groove (**Figure 3.10A**).

Table 3.6: Potential Energy (kcal/mol) of SJG-136 covalently bound to every potential reacting guanine of NF-κB-1, NF-κB-2, EGR-1, AP-1 and STAT3 sequences.
(F) indicates a forward adduct (*i.e.*, towards the TTT-loop) and (R) indicates a reverse adduct (*i.e.*, away from the DNA loop).

Sequence	Adduct Type	Reacting Guanine	Potential Energy (kcal/mol) (covalent)	Ranking
NF-κB-1	**Mono-adducts**	G3 (F)	-4561.61	1
		G2 (F)	-4560.78	2
		G7 (F)	-4560.14	3
	Cross-Links	G2-G19	-4543.53	1
		G1-G3	-4541.22	2
		G3-G19	-4537.75	3
NF-κB-2	**Mono-adducts**	G5 (F)	-4737.04	1
		G17 (R)	-4736.88	2
		G4 (F)	-4734.72	3
	Cross-Links	G1-G3	-4718.88	1
		G4-G17	-4716.83	2
		G2-G4	-4712.32	3
EGR-1	**Mono-adducts**	G6 (R)	-4423.57	1
		G19 (F)	-4422.44	2
		G7 (R)	-4421.78	3
	Cross-Links	G5-G14	-4420.56	1
		G3-G5	-4413.51	2
		G5-G19	-4409.82	3
AP-1	**Mono-adducts**	G7 (R)	-3230.85	1
		G17 (F)	-3229.70	2
		G1 (F)	-3227.69	3
	Cross-Links	G1-G17	-3212.85	1
		G7-G17	-3200.44	2
STAT3	**Mono-adducts**	G7 (R)	-3708.78	1
		G7 (F)	-3706.57	2
		G19 (F)	-3704.43	3
	Cross-Links	G6-G19	-3687.29	1
		G6-G18	-3681.42	2
		G5-G7	-3680.12	3

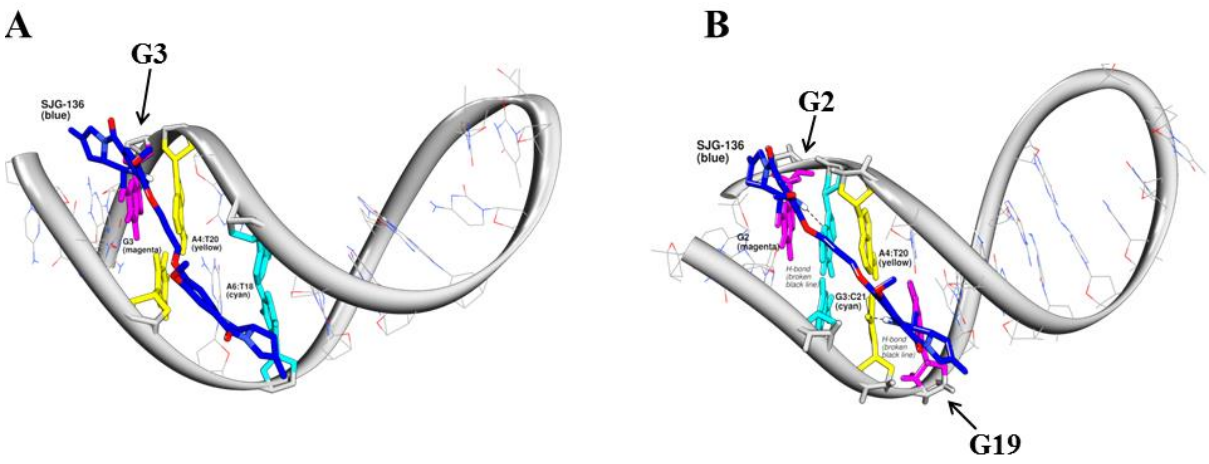


Figure 3.10: Low energy snapshots: **A**, SJG-136 (blue) covalently bound to G3 (magenta) of NF-κB-1. The central methylene linker of SJG-136 forms extensive van der Waals interactions with the A4:T20 base pair (yellow) and an unreacted PBD of the PBD dimer forms non-covalent interactions with the A6:T18 base-pair (cyan), allowing the molecule to bind snugly in the DNA minor groove. **B**, SJG-136 (blue) covalently bound to G2 and G19 (magenta) of NF-κB-1. The central methylene linker of SJG-136 forms extensive van der Waals interactions with the A4:T20 base pair (yellow) with stabilising hydrogen bonds between the N10-proton of one PBD and the ring nitrogen (N3) of the adjacent G3, and between the N10-proton of the second PBD and O4 of the neighbouring T20.

The most energetically preferred cross-linked adduct is the G2-G19 adduct, where SJG-136 spans six base-pairs across the sequence 5'-Pu-GGAC-Pu-3' (**Table 3.6**). In a similar manner to results from NF-κB-2 sequence molecular dynamics simulations, the 5'-Pu-GGAC-Pu-3' adduct is very similar to the most preferred adduct of SJG-136 5'-Pu-GATC-Py-3'. Low energy snapshots of the simulation suggest that SJG-136 embeds in the minor groove due to the formation of non-covalent interactions between the central methylene linker and A4:T20 (**Figure 3.11A**). Hydrogen bonds between N10-H of the PBDs and adjoining bases (*i.e.*, G3 and T20) also assist in restraining the DNA in the minor groove. Similarly, as both PBDs are in an ideal location for covalent attack of G2 and G19, the DNA cross-link formed induces little distortion of the DNA base-pairing of the sequence, and the G2-G19 cross-link is therefore the most preferred adduct. Other potential cross-links in the NF-κB-1 sequence include three base-pair (*e.g.*, G1-G3) and five base-pair (*e.g.*, G3-G17) sequences, which are known to be less preferred than a central four base-pair sequence such as 5'-GATC-3'. As a result, these potential energy calculations agree with previously published biophysical studies[150].

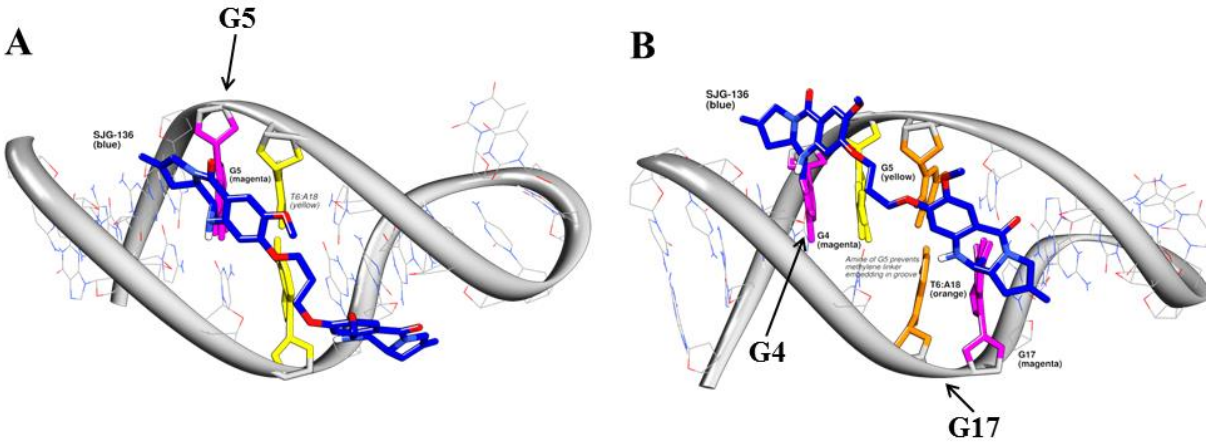


Figure 3.11: Low energy snapshots: **A**, SJG-136 (green) covalently bound to G5 (magenta) of NF-κB-2 sequence in reverse orientation (*i.e.*, pointing away from the TTT-loop). **B**, SJG-136 (blue) covalently bound to G4 and G17 (magenta) of NF-κB-2. G5 (which impedes ligand binding) is yellow and bases involved in non-covalent interactions are illustrated in orange sticks.

Similar molecular dynamics simulations were undertaken on NF-κB-2 sequence to predict the most favoured reacting guanine. As SJG-136 is known to form mono-adducts and both inter- and intrastrand cross-links[96], study design focused on the identification of sequences within the NF-κB-2 consensus site which could potentially form cross-linked or mono-alkylated SJG-136 adducts. Specific criteria were set in study design whereby cross-linked (both intra- and inter-) adducts were required to be a minimum of three base-pairs in length, in line with the previous literature[205]. In the case of mono-alkylated adducts, a base-pair span of at least four base-pairs was considered necessary for reactions. As such, adducts such as the mono-alkylated G3 adduct in forward (*i.e.*, loop-facing) orientation were considered viable, whereas adducts such as G2 in reverse (pointing away from the loop) were not. Similarly, adducts such as a loop-facing G14 or G15 mono-adduct were not assessed due to steric hindrance introduced due to the presence of the TTT-loop[206]. In all, nine potential mono-alkylated adducts and fourteen potential cross-links were identified and the three most preferred are outlined in **Table 3.6**.

Potential energy calculations (kcal/mol) averaged over the duration of each simulation suggest the most likely mono-alkylated adduct is a G5 adduct where SJG-136 is in forward orientation (*i.e.*, pointing towards the loop) (**Table 3.6**). In this instance, it is

likely that stabilising van der Waals interactions promote the interaction of the methylene chain of the ligand with the T6:A18 base-pair, assisting ligand accommodation in the minor groove of DNA (**Figure 3.11B**).

Potential energy calculations (kcal/mol) suggest that the G1-G3 is the most stable cross-link, and, therefore, most likely to occur. This observation is unusual, given prior evidence which suggests PBDs require a three base-pair site to bind (*e.g.*, Pu-G-Pu)[125]. Visual analysis of the molecular dynamics simulation results show that although base-pairing is maintained throughout the simulation, the ligand is highly distorted. As such, the G1-G3 adduct (and therefore alkylation of G1) is unlikely to preferentially occur. The second most-preferred adduct (based on potential energy calculations) is a G4-G17 adduct (**Figure 3.11A**). SJG-136 is known to preferentially bind in an interstrand manner across a six base-pair sequence (Pu-GATC-Py) and as the G4-G17 adduct is remarkably similar in content, *i.e.*, Pu-GGTC-Pu, it is likely that this is the preferred cross-linked adduct. This assumption is further supported by the fact that SJG-136 was shown through biophysical experiments to bind to examples of 5'-Pu-GXXC-Py-3' (where X is A, T, G or C)[207]. Moreover, although the exocyclic amine of G6 protrudes into the minor groove preventing the formation of adduct stabilising hydrogen bonds and full ligand accommodation, there is little ligand distortion evident in simulations of SJG-136 covalently bound to 5'-GGTC(G)-3' (**Figure 3.11A**), and as such it is likely that a G4-G17 adduct is most preferred.

In the case of EGR-1 sequence, SJG-136 is known to form a potential G1-G3 cross-link across a three base-pair sequence (in this case 5'-GCG-3'), thus, this adduct was included in the analysis. Similarly, SJG-136 is known to be capable of forming an extended five base-pair cross-link, as such the G1-G5 cross-link was also investigated. Adducts such as a loop-facing G9 mono-adduct or a loop-facing G14 mono-adduct were not assessed due to steric hindrance introduced by the presence of the TTT-loop[208]. In all, eleven potential mono-alkylated adducts and ten potential cross-links were identified and the three most preferred are shown in **Table 3.6**.

Potential energy calculations (kcal/mol) averaged over the duration of each simulation suggest that the most likely mono-alkylated adduct is a G6 adduct where SJG-136 is in reverse orientation (*i.e.*, pointing away from the loop) (**Figure 3.12A**). The alkylation of

G6 by a single PBD of the dimer SJG-136 allows the molecule to maximise its interaction with DNA through van der Waals interactions with GC base-pairs (orange bases, **Figure 3.12A**). The adduct is also stabilised by the formation of a hydrogen bond between the NH of the PBD and O2 of C17. Similarly, for the second and third most likely mono-alkylated adducts (G19 in forward orientation and G7 in reverse orientation), the dimer is positioned in an identical orientation, maximising interactions with DNA in the process.

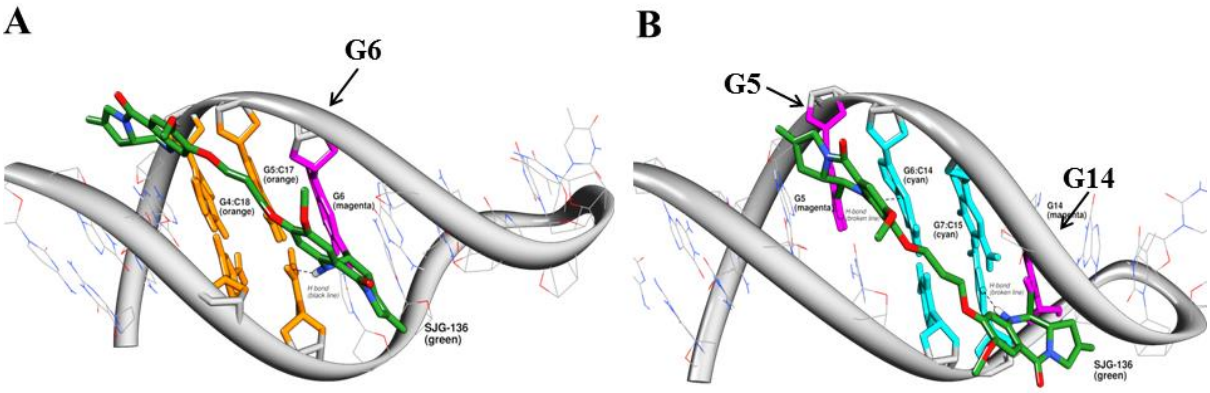


Figure 3.12: Low energy snapshots: **A**, SJG-136 (green) covalently bound to G6 (magenta) of EGR-1 sequence in reverse orientation (*i.e.*, pointing away from the TTT-loop). Bases involved in non-covalent interactions are illustrated in orange sticks. **B**, SJG-136 (green) covalently bound to G5 and G14 (magenta) of EGR-1 sequence in reverse orientation (*i.e.*, pointing away from the TTT-loop). Bases involved in non-covalent interactions are illustrated in cyan sticks.

The most likely cross-linked adduct (according to potential energy calculations) is the G5-G14 interstrand adduct (**Table 3.6**, **Figure 3.12B**). SJG-136 is known to be highly interactive with the sequence 5'-Pu-GATC-Py-3'[94], and is also known to form other covalent attachments to DNA, including extended cross-links such as 5'-Pu-GAATC-Py-3'[205] and examples of 5'-Pu-GXXC-Py-3' (where X is A, T, G or C)[207]. As the G5-G14 adduct 5'-Pu-GGGC-Pu-3' is an equivalent span and content to other preferred interacting sequences, the preference for formation of a G5-G14 adduct would be in accord with previous observations. Crucial to the strong interaction of SJG-136 with 5'-GGGC-3' is the accommodation of the molecule in the minor groove and the formation of two

hydrogen bonds; the first between the NH of the PBD bound to G5 and the N3 of G6 and the second between the NH of the PBD bound to G14 and O2 of C15 (**Figure 3.12B**).

Molecular dynamics simulations were conducted on SJG-136 covalently bound to every potential reacting guanine in the AP-1 consensus sequence (*i.e.* G1, G7, G12, G17) in order to rationalise the formation of one adduct peak in HPLC studies. During study design, potential reaction sites were chosen based on the span of the molecule. For example in the case of G1, a terminal base, an adduct in reverse orientation (*i.e.*, pointing away from the loop) was not selected for analysis due to the known preference of a PBD for a three base-pair motif for interaction. Similarly, for G7 and G12, a forward adduct (*i.e.*, pointing towards the loop) was not analysed due to the potential steric hindrance associated with the TTT-loop, however G17 in both orientations was investigated due to the possibility of adduct formation in both orientations.

Although SJG-136 is known to prefer to form adducts over a six or seven base-pair sequence (*i.e.*, Pu-GATC-Py, Pu-GATG-Py, Pu-GAATG-Py or Pu-GAATC-Py[205]), potential energy calculations suggest that the adduct spanning 5'-GAC-3' (*i.e.*, G1-G17) is preferred to the extended adduct (5'-CATTG-3'). Although surprising due to the fact that the extended 5'-ATT-3' base sequence in the centre of the DNA sequence should form extensive van der Waals interactions with the methylene linker of the ligand, this is in fact in accord with previous studies on PBD molecules, which suggest A-ring 3'-orientation (as evident in 5'-GAC-3') is preferred to A-ring 5'-orientation (5'-CATTG-3'). It is also interesting to note that although 5'-GAC-3' (*i.e.*, G1-G17, **Figure 3.13A**) is the most preferred cross-linked adduct, potential energy calculations suggest that a mono-alkylated adduct in the same orientation (G17 facing away from the loop) may also form. In HPLC studies, a single adduct was observed to form and potential energy calculations suggest that if a mono-adduct forms, this is likely to be G7 (**Figure 3.13B**).

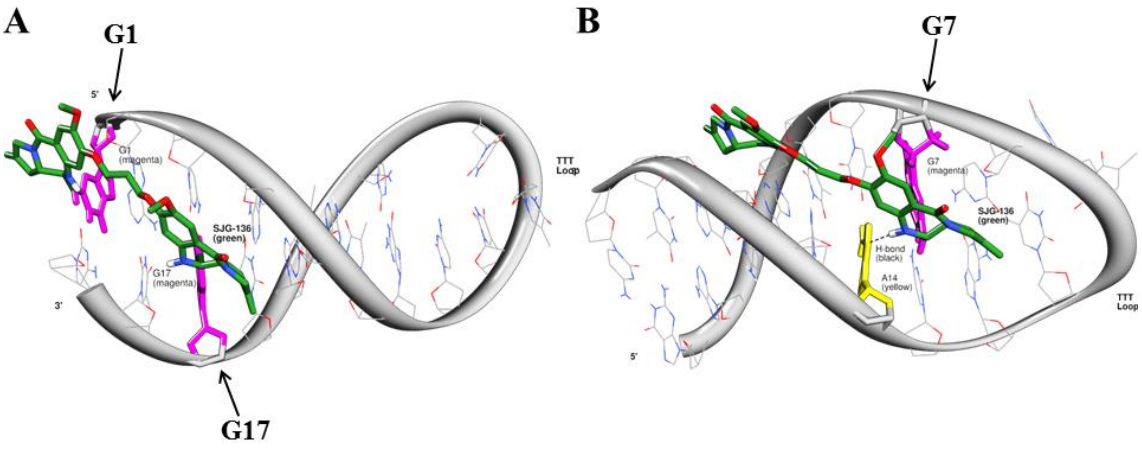


Figure 3.13: Low energy snapshots: **A**, SJG-136 (green) covalently bound to G1 and G17 (magenta) of AP-1 sequence forming an interstrand cross-link across the sequence 5'-GAC-3'. **B**, SJG-136 (green) covalently bound to G7 of AP-1 sequence. The NH group of the PBD forms a hydrogen bond with an adjacent adenine base, A14 (yellow).

Potential energy calculations (kcal/mol) conducted on the STAT3 sequence suggests a definite preference for the G7 mono-adduct for SJG-136 in this sequence (**Table 3.6**, **Figure 3.14B**). During interaction with its favoured sequence Pu-GATC-Py, SJG-136 is thought to form van der Waals interactions between the methylene linker and A:T base-pairs, while concurrently forming sequence-selective hydrogen bonds with bases adjoining the covalently modified guanines[209].

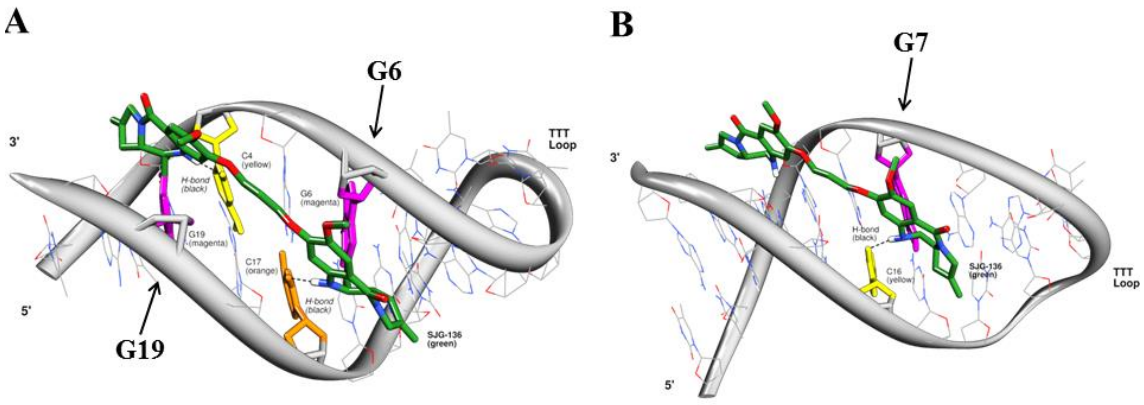


Figure 3.14: Low energy snapshots: **A**, SJG-136 (green) covalently bound to G6 and G19 (magenta) of the STAT3 sequence. The NH groups of each PBD form hydrogen bonds to cytosine residues (*i.e.*, C17 illustrated in orange and C4 illustrated in yellow) and form covalent bonds with G6 and G19, spanning the sequence 5'-Py-CCGG-Pu-3'. **B**, SJG-136 (green) covalently bound to G7 (magenta) of the STAT3 sequence. The NH group of the PBD forms a stabilising H-bond to C16 (yellow).

In the case of the STAT3 sequence, every potential cross-linked adduct must form around the central 5'-CCGGG-3' tract. This would pose considerable steric difficulties due to the presence of C2-amino groups of guanines beneath the methylene linker in the case of every cross-linked adduct, thus preventing the interaction of SJG-136 with the minor groove floor. According to potential energy calculations, the G6-G19 adduct is the most favoured (potential energy of -3687.29 kcal/mol) of the cross-linked adducts as it would span across six base-pairs (*i.e.*, Py-CCGG-Pu) forming hydrogen bonds in the process (**Table 3.6**, **Figure 3.14A**), followed by G6-G18 (Py-CGG-Pu), G5-G7 (Py-GGG-Pu), G7-G18 (Py-CGGG-Pu) and G7-G19 (Py-CCGGG-Pu). It is relevant to mention that the presence of a guanine triplet on the forward strand (5'-GGG-3') appears to interfere with the binding of SJG-136, possibly occurring due to steric hindrance between the guanine residues and methylene linker of SJG-136.

The top three most favoured guanine residues and orientations are illustrated (**Table 3.6**) and it is evident that SJG-136 requires a considerable minor groove environment for interaction with DNA. For example, G7, G6 and G19 all feature in the top five most favoured drug/DNA adducts, and each of these bases is positioned at the centre of the DNA sequence, providing a large minor groove environment for ligand interaction and

thus maximising non-covalent interaction. G19, on the other hand, is located towards the 3'-end of the DNA and produces poor potential energy results when facing away from the TTT-loop, and is therefore less favoured.

Molecular dynamics results, overall, suggest that DNA cross-links are less favoured than mono-adducts in the case of the STAT3 consensus sequence, and it is likely that the three adducts produced in HPLC studies are mono-adducts. Furthermore, based on potential energy calculations, molecular modeling studies indicate that adducts should form in the priority: G7 (adduct facing away from loop) = G7 (loop-facing) > G19 (loop-facing) > G6 (loop-facing) = G6 (adduct facing away from loop), and, therefore, it is expected that adducts involving each of the three DNA bases G7, G19 and G6 are the most likely adducts formed.

3.5.2 Summary of molecular modeling study

Molecular dynamics simulations and potential energy calculations indicate that within the NF-κB-1 consensus sequence the most likely mono-adduct is formed at the guanine in position G3 (pointing towards the loop) as non-covalent interactions between the central methylene linker of the dimer and the A4:T20 base-pair assist the ligand to accommodate within the DNA-minor groove. The interstrand cross-link at G2-G19 is the most preferred cross-link adduct due to the formation of non-covalent interactions between the central methylene linker and A4:T20. Additionally, hydrogen bonds between N10-H of the PBDs and adjoining bases (*i.e.*, G3 and T20) also assist in restraining the DNA in the minor groove.

In the case of NF-κB-2 sequence, the molecular modeling study suggested that the most likely mono-alkylated adduct to occur is at the guanine in position G5 in forward orientation (*i.e.*, pointing into the loop) due to stabilising van der Waals interactions between the methylene chain of SJG-136 with the T6:A18 base-pair of the DNA. Based on the potential energy calculations the most stable cross-link is the intrastrand cross-link at G1 and G3.

For EGR-1 sequence, molecular dynamics simulations and potential energy calculations showed that the most likely mono-alkylated adduct is formed at G6 in reverse orientation

(*i.e.*, pointing away from the loop) due to the maximum interaction of SJG-136 with the EGR-1 sequence though van der Waals interactions and hydrogen bond formation. According to the potential energy calculations the most likely cross-linked adduct is the interstrand cross-link at G5 and G14 due to the strong interaction of SJG-136 with 5'-GGGC-3'.

Molecular dynamics simulations and potential energy calculations conducted on AP-1 sequence suggested that the most likely mono-alkylated adduct is formed at G7 in reverse orientation (*i.e.*, away from the DNA loop) where the NH group of the PBD forms a hydrogen bond with an adjacent adenine base, A14. Based on the potential energy calculations the most likely cross-link adduct is the shorter interstrand cross-link at G1 and G17 which suggests that A-ring 3'-orientation (as evident in 5'-GAC-3') is preferred to A-ring 5'-orientation (5'-CATTG-3').

Finally, the molecular modelling study carried out on the STAT3 consensus sequence indicated that mono-alkylated adduct formation is preferred over cross-links. The most preferred mono-adduct may be formed at G7 orientated away from the loop (*i.e.*, reverse). Based on the data, it is anticipated that the most favoured cross-linked adduct may be formed between G6 and G19.

3.5 Conclusion

In conclusion, this is the first report of the PBD dimer SJG-136 binding to the consensus sequences of the oncogenic transcription factors NF-κB, EGR-1, AP-1 and STAT3. The HPLC/MS study has suggested that SJG-136 may be capable of binding to the cognate sequences of NF-κB, EGR-1, AP-1 and STAT3. Interestingly, the results revealed significant differences in the extent and rate of adduct formation between the individual sequences which may explain the previously observed differences in activity of SJG-136 in various human tumours. For example, SJG-136 appeared to form rapidly adduct with NF-κB-2 and AP-1 sequences while it reacted moderately with the EGR-1 sequence and slowly with NF-κB-1 and STAT3 sequences, despite their comparable GC-contents. Furthermore, significant variations in the rate of adduct formation were observed. As a dimer, SJG-136 can form intra-, inter-strand and mono-alkylated adducts within the DNA

minor groove[96, 136]. Even more important, on the basis of the HPLC/MS data, SJG-136 formed three distinct adducts with the STAT3 sequence, two different adducts with the NF-κB-1 sequence, while only single adduct formation was observed in the case of NF-κB-2, EGR-1 and AP-1 sequences. These results indicate that SJG-136 favours one particular binding site within the NF-κB-2, EGR-1 and AP-1 sequences, as only one particular type of adduct had been formed. Moreover, the most remarkable observations have been made on the analysis of the AP-1 sequence. Based on the obtained data, the ligand has demonstrated a preference for this particular sequence as it reacted rapidly and to the full extent after 5 min with one main adduct formed. However, the limitations of the developed technique have been discussed previously. Therefore, the appeared high reactivity cannot be confirmed using this method and future studies have to be carried out in order to support this observation. The results were supported to some extent by FRET, CD and molecular modeling studies. Important to note are the different drug and DNA concentrations used for HPLC/MS (DNA 25 μM, SJG-136 100 μM), FRET (DNA 400 nM, SJG-136 2 μM) and CD (DNA 5 μM, SJG-136 μM) studies. All these concentrations have been chosen according to previous studies carried out on PBD monomers and dimers with short oligonucleotides[91, 96, 210]. However, although different concentrations for the DNA and SJG-136 were used during this study, a 4:1 ratio (SJG-136/DNA) was kept for the HPLC, MALDI-TOF and CD study which was previously reported to ensure completion of adduct formation within a reasonable time frame[210a]. Due to the different concentrations used, the results may not be directly comparable. During the FRET study, the concentration of SJG-136 was gradually increased as no differences in ΔT_m could be observed for most of the sequences (apart from AP-1).

The data presented in this report suggests that SJG-136 may exert its pharmacological effect through interaction with cognate sequences of oncogenic transcription factors in addition to previously reported mechanisms including; strand breakage, inhibition of enzymes including endonucleases and RNA polymerases and arrest of the replication fork[88a, 203]. This wealth of new data considerably adds to the knowledge of the mechanism of action of SJG-136 and is of utmost importance for the correct interpretation of the results of biochemical and pharmacological assays.

To investigate whether a correlation exists between the biophysical results and the behaviour of SJG-136 *in vitro*, the ligand was evaluated in a biological study utilising two human tumour cell lines. Based on the surprising results obtained for the AP-1 sequence, the colon carcinoma cell line HT-29 has been selected for the experiments due to the reported involvement of AP-1 in the development and progression of colorectal tumorigenesis[54, 211]. Similarly, the STAT3-dependent breast cancer cell line MDA-MB-231 was chosen for the biological evaluation of SJG-136 as the ligand has demonstrated a low preference for STAT3 during the HPLC/MS study. Elevated levels of STAT3 have been detected in breast cancer and are presumed to promote tumour growth and metastasis[64, 212]. Another reason for the selection of AP-1 and STAT3 was their intracellular localisation. STAT3 is present in a latent state in the cytosol and enters the nucleus after its activation through phosphorylation where it binds to promoter and enhancer regions on the DNA[63a], whereas AP-1 is located within the nucleus where it can rapidly bind to its consensus sequences on the DNA[53b]. The results of the biological study are presented in the following chapter (Chapter 4).

Chapter 4: Mechanistic confirmation of transcription factor inhibition by the PBD dimer SJG-136

4.1 Background

Proteins are the major components of cells in multicellular organisms. They perform almost all cellular functions in the human body. Enzymes, such as proteases, synthases and kinases catalyse and promote the chemical reactions in cells[213]. The transition of small molecules into the extracellular environment is carried out by proteins that are integrated into the membrane and form channels and pumps (*i.e.*, Ca^{2+} pump in muscle cells)[213]. Cell-to-cell messages and signals, as well as intracellular communication between the cytosol and the nucleus, are mediated by proteins[213]. Further examples of proteins are antibodies which bind with high specificity to specific proteins on the membrane of bacteria and viruses in order to label them for autologous destruction[213].

Many proteins are constantly required due to their indispensable tasks and functions in multicellular organisms (*e.g.*, structural proteins of chromosomes)[214]. Those proteins are constitutively expressed in order to maintain their constant concentration in the organism. Other proteins are present only in specialised cells where they perform particular functions (*i.e.*, haemoglobin in red blood cells) or are produced in response to external signals and changes in the environment of cells such as signalling or regulatory proteins[214]. The process from transcription to translation is exceedingly complex and involves various steps. It will be described only briefly in this chapter.

Transcription factors (TFs) are sequence-specific DNA binding proteins. They contain DNA binding domains which enable them to bind to particular sequences of DNA called enhancer or promoter sequences. The DNA promoter sequence can be located near the transcription start site. TFs can also bind to regulatory sequences such as enhancer sequences located thousands of nucleotide pairs upstream or downstream away from the gene being transcribed. This step recruits RNA polymerase II to the promoter sequence (a short DNA sequence which indicates the transcription start-point for the enzyme) and where a transcription initiation complex is formed at the template strand consisting of transcription factors and RNA polymerase II. Following this, the RNA polymerase II

begins to translate the encoded DNA sequence and synthesise a complementary RNA single strand using the substrates ATP, CTP, UTP and GTP. When the RNA strand has been elongated for about 25 nucleotides the 5'-end of the emerging RNA is modified by the addition of a methyl cap. This modification signals the 5'-end of the precursor-mRNA and assists the cell to differentiate mRNAs from other existing types of RNAs in the cell. However, eukaryotic genes are not composed of continuously protein-coding DNA sequences (exons). Moreover, they are interspersed with longer non-coding regions, also referred to as introns. The process of intron removal and joining the exons to form one consistent protein-coding mRNA strand is termed RNA splicing. RNA splicing contributes to the genomic diversity and tissue specificity. This process allows the cell to rearrange exons, resulting in the production of a significant number of protein isoforms from a single gene[215]. Moreover, mutations can easily accumulate in the genome without causing problems due to the bigger fraction of non-coding sequences. Therefore, ligation of the exons decreases the likelihood of mutations to occur.

Once transcription has been terminated, the 3'-end of the nascent mRNA molecule is modified by addition of about 200 adenine nucleotides (polyadenylation). Only when modifications, such as 5'-end capping, RNA splicing and 3' polyadenylation have all been completed, the mature mRNA is transported from the nucleus into the cytosol where the nucleotide sequence of a gene is translated into the amino acid sequence of a protein. This conversion from mRNA into a particular protein sequence is also known as translation. Translation is initiated by the binding of a small ribosomal subunit at AUG start codons which in turn is recognised by a special initiator tRNA molecule. Following this, a large ribosomal subunit assembles at the small subunit and forms the complete ribosome (located on the rough endoplasmic reticulum) in which protein synthesis takes place. Aminoacyl tRNAs that carry specific amino acids bind with their anticodon to the complementary codon on an mRNA molecule. Appropriate aminoacyl-tRNA synthetases recognise codons for their specific amino acid and connect them by forming a peptide bond between the carboxyl group at the end of a growing polypeptide chain and the free NH_2 group of an incoming amino acid. Polypeptide chains are so produced stepwise from their N-terminus towards their C-terminus. Stop codons containing sequences UAA, UAG and UGA on mRNA mediate translation termination to the ribosomes. As a consequence, a water molecule is attached to the peptidyl-tRNA instead of an amino acid which locks the carboxyl group. The synthesised proteins then undergo various post-

translational modifications such as quality control checking which all occurs in the Golgi network. Packed into vesicles, they are transported throughout the cell and distributed depending upon their localisation signals (if present) the synthesised protein is released into the cytosol. Lastly, the newly synthesised polypeptide chain is folded into its unique and correct three-dimensional structure with the assistance of chaperones. Misfolding of proteins will lead to their malfunction and cause damage to the cells. Therefore, they are destroyed by proteolytic degradation initiated by ubiquitin *via* the proteosome.

It has become increasingly clear that each step between transcription initiation and translation has to be strictly regulated in order to avoid the synthesis of incorrect functioning proteins that will have crucial consequences for the organism they are expressed in. In the context of the aims of this project, only the post-transcriptional controls which take place following gene expression will be discussed.

The earliest stage of post-transcriptional regulation is premature transcription attenuation. Anti-attenuators (special RNA binding proteins) assemble at nascent pre-mRNA and allow RNA polymerase II to transcribe a gene continuously without stopping[216]. Another post-transcriptional control mechanism is RNA editing. During this process, the nucleotide sequence of the mRNA transcript is altered by insertion or deletion of nucleotides or exchange of one base to another leading to various protein products[217]. Exosomes (large protein complexes with exonuclease activity) degrade any incompletely processed or damaged RNAs before they are exported to the cytosol[214a]. Additionally, translation initiation can be negatively regulated by translational repressors that bind to the 5'-end of the mRNA and prevent the small ribosomal subunit recognising the start codon AUG and thus preventing the initiation of translation[214b]. A further substantial control in protein synthesis is the phosphorylation of eIF2, a eukaryotic initiation factor, by kinases in response to stress situations, including deprivation of growth factors and nutrients or infection by viruses with the aim of decreasing cellular damage or inducing apoptosis[218]. This process inhibits the binding of methionyl initiator tRNA to the small ribosomal subunit. Alternatively, translation can be initiated at other specialised positions on the mRNA apart from the first start codon AUG at the 5'-end. Those internal recognition sites are called 'internal ribosome entry sites' (IRESs) and can be activated without the need of a 5'-cap structure[219]. The benefit of this regulation process for the cell is that translation of special mRNAs can take place even in the presence of a general

decrease in protein synthesis initiation. Many eukaryotic mRNAs are unstable, with half-lives of 30 minutes or even less[214a]. Consequently, translation of those unstable mRNAs is regulated by changing their stability in the cytosol. Shortening of the poly-A tail by exonucleases or its cleavage at once by specific endonucleases leads to rapid mRNA degradation[220]. On the other hand, lengthening of the poly-A tail increases the stability of mRNAs. Having discussed the degradation of aberrant transcribed RNAs in the nucleus by exosomes an additional translation control mechanism occurs in the cytosol also referred to as 'nonsense-mediated mRNA decay'[221]. This process aims to degrade any aberrant mRNAs before efficient translation into proteins resulting in their malfunction. It has become increasingly clear that the exceedingly complex network of post-transcriptional control plays a crucial and substantial role in the translation of mRNA into protein. But regardless of the versatility of all control processes, their fundamental role is to prevent translation of any incorrectly transcribed mRNAs into malfunctioning proteins which cause significant damage to cells.

In context to the body of this work, interaction of the small-molecules with consensus sequences of TFs results in inhibition of transcription and all the following process as described in detail above. This may lead to inhibition of proteins that are known to play a major role in cancer development and progression. Therefore, targeting the consensus sequences of TFs is an option to treat malignancies, and is an interesting potential mechanism of action of SJG-136.

4.2 Objectives

Biophysical studies carried out on the cognate sequences of the oncogenic TFs NF-κB, EGR-1, AP-1 and STAT3 demonstrated that the PBD dimer SJG-136 may exert its pharmacological effect also through TF inhibition, in addition to previously reported mechanisms including strand breakage, inhibition of enzymes including endonucleases and RNA polymerases, and arrest of replication fork[88a, 203] by binding to the consensus DNA target sequences (Chapter 3). During the HPLC/MALDI-TOF-MS study adduct formation between SJG-136 and NF-κB, EGR-1, AP-1 and STAT3 sequences was observed and identified, showing the ability of the ligand to interact with these particular TF binding sites. The HPLC/MS results were supported by FRET, CD, molecular dynamics simulations and free energy binding calculations where SJG-136 demonstrated the capability to interact with all studied target DNA sequences.

As a consequence of the data obtained from the biophysical study we aimed to confirm the ability of SJG-136 to inhibit TF-mediated processes, such as gene and consequently protein expression *in vitro* by investigating the effect of SJG-136 on the human tumour breast cancer cell line MDA-MB-231 and the human colon carcinoma cell line HT-29. The hypothesis was that if the ligand is capable of binding to consensus sequences of the oncogenic transcription factors STAT3 and AP-1, the expression of genes that require STAT3 or AP-1 activation was assumed to decrease. As a consequence, subsequent protein synthesis of incorrectly regulated or functioning proteins involved in cancer development and progression was expected to be down-regulated. To address this issue, two most frequently used and established methods were applied in this study. Polymerase chain reaction was utilised to investigate changes in gene expression profiles and quantitative Western blotting was used to analyse the effect of SJG-136 on protein expression.

4.3 Selection of human tumour cell lines for biological study

4.3.1 Human breast cancer cell line MDA-MB-231

The MDA-MB-231 cell line (MDA-MB-231 ATCC® HTB-26™) is a triple negative metastatic human breast cancer cell line originally isolated in the early 1970s[222]. It is an immortal cell line due to the presence of an active telomerase, which prevents telomere shortening and allows the permanent division of cells[223]. Cells possess a spindle shaped morphology[224] as illustrated in **Figure 4.1A** and **B**. General information about MDA-MB-231 cell line obtained from ATCC® is shown in **Table 4.1**. Extensive research has been undertaken in order to identify mechanisms responsible for breast cancer cell migration, metastasis and its poor prognosis and STAT3 has been shown to promote tumour growth as well as tumour metastasis due to its elevated levels in breast cancer[64, 212, 225]. This particular cell line has been selected for our biological experiments as SJG-136 showed the lowest preference for the cognate sequence of the transcription factor STAT3 during the biophysical study. Very slow reaction was observed with the STAT3 consensus sequence, with only ~14% adducts formed after 24 hours incubation of SJG-136 with DNA. Furthermore, multiple adducts were formed. All these data suggest that SJG-136 is less selective for STAT3 with no preferred binding site. This cancer cell line did not act as a negative control as it was chosen to investigate whether the differences in activity observed for STAT3 and AP-1 during the biophysical study correlate with the *in vitro* behaviour of SJG-136 toward STAT3 and AP-1 dependent human tumour cell lines.

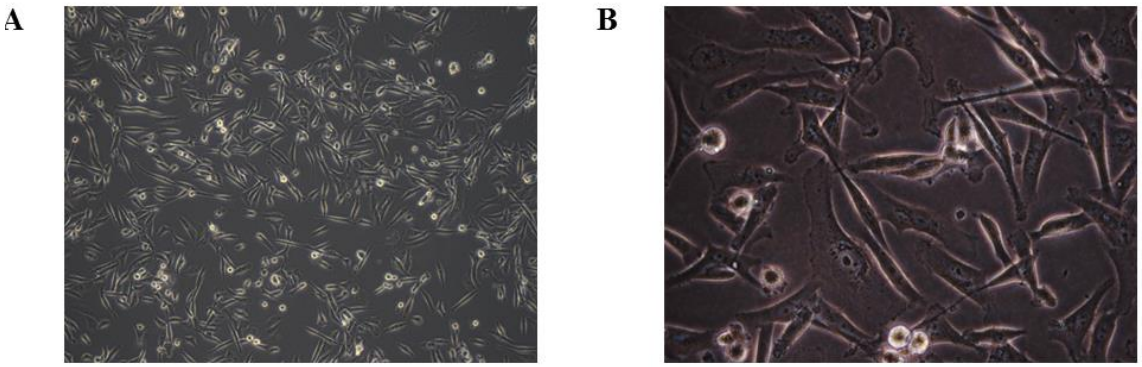


Figure 4.1: Microscope image of human breast cancer cell line MDA-MB-231 used in this study. **A**: 20 x magnification; **B**: 40 x magnification. Passage number: 20.

Table 4.1: General information about MDA-MB-231 ATCC® HTB-26™.

Organism	*Homo sapiens*
Tissue	mammary gland/breast; derived from metastatic site: pleural effusion
Cell Type/Morphology	epithelial
Culture Properties	adherent
Disease	adenocarcinoma
Age	51 years adult
Gender	Female
Ethnicity	Caucasian

4.3.2 Human colon carcinoma cell line HT-29

The HT-29 cell line (HT-29 ATCC® HTB-38™) is a human colorectal adenocarcinoma cell line with epithelial morphology. This cell line was established in 1964. Under standard culture conditions, these cells grow as a nonpolarised, undifferentiated multilayer[226] (**Figure 4.2A** and **B**). General information about the cell line HT-29 obtained from ATCC® is shown in **Table 4.2**. Studies have been carried out to investigate the involvement of the TF AP-1 in human colorectal tumour promotion and progression[54, 227]. The results have demonstrated over-expression of crucial AP-1 family members (*i.e.*, Jun and Fos), thus contributing and promoting colorectal tumourigenesis. The cell line HT-29[54, 211, 227] was chosen for our biological evaluation of SJG-136 on the basis of the surprising observations made on the AP-1 cognate sequence during the biophysical study. Remarkably rapid adduct formation was observed between SJG-136 and sequence AP-1 with ~100% adduct formed after 5 min incubation. In addition, only one particular adduct type was formed suggesting a favoured binding site of SJG-136 within sequence AP-1.

A **B**

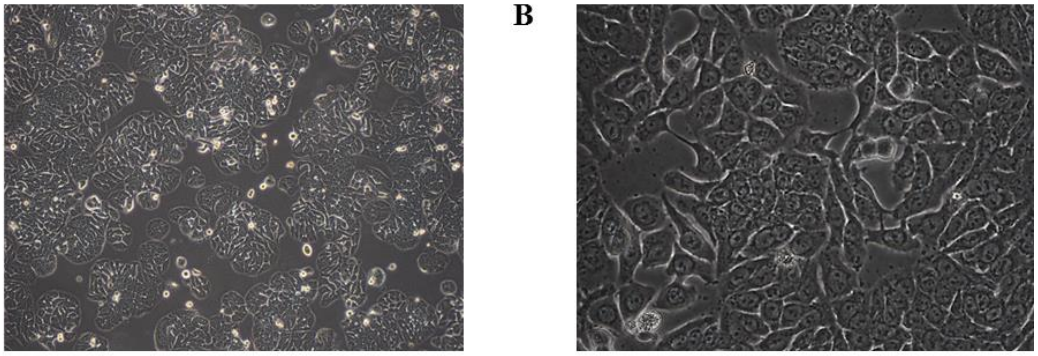


Figure 4.2: Microscope image of human colon carcinoma cell line HT-29 used in this study. **A**: 20 x magnification; **B**: 40 x magnification. Passage number: 16.

Table 4.2: General information about HT-29.

Organism	*Homo sapiens*
Tissue	Colon
Cell Type/Morphology	epithelial
Culture Properties	Adherent
Disease	colorectal adenocarcinoma
Age	41 years adult
Gender	Female
Ethnicity	Caucasian

4.4 Genes investigated in PCR study on MDA-MB-231 and HT-29

4.4.1 *Bcl-2*

Bcl-2, also known as B-cell like lymphoma 2, is a member of the Bcl-2 family of proteins that regulate all major types of cell death, including apoptosis, necrosis and autophagy[228]. The Bcl-2 protein has been shown to play a central role in cell death regulation[229]. Bcl-2 was discovered in Non-Hodgkin's lymphomas in 1985[230]. Elevated levels of Bcl-2 expression have been documented in most chronic lymphocytic leukaemia due to mutations in encoding genes[231]. Aberrant Bcl-2 gene expression results in dysregulation of crucial cell-death mechanisms and contributes to the pathogenesis and progression of human cancers[228].

4.4.2 *Cyclin D1*

The progression from one cell cycle stage to another is regulated by regulatory proteins, called cyclin-dependent-kinases (CDKs), which in turn are activated by other proteins, the cyclins[232]. Cyclin D1 binds to CDK4 and CDK6 in early G_1 phase during which the cell is preparing for DNA synthesis[232]. The consequences of cyclin D1 overexpression are disruption and incorrect cell cycle regulation. Aberrant cyclin D1 expression is associated with the development and progression of several human cancers, such as breast, oesophagus, bladder and lung cancer[233].

4.4.3 *STAT3*

Signal transducer and activator of transcription 3 (STAT3) belongs to the STAT family composed of seven proteins (STATs 1, 2, 3, 4, 5a, 5b, and 6) that transfer external signals to the nucleus in order to regulate gene transcription[64]. STAT3 regulates the expression of genes that are crucial for cell differentiation, proliferation, apoptosis, angiogenesis, metastasis, and immune responses[65a]. STAT3 is considered as an oncogene as in many human malignancies constitutive STAT3 activation and consequently elevated levels of STAT3-dependent genes has been associated with defects in apoptosis regulation, tumour cell proliferation and survival, angiogenesis and metastasis[65a, 234].

4.4.4 *STAT1*

In contrast to STAT3, STAT1 is a tumour suppressor[235] as it negatively regulates tumour angiogenesis, growth and metastasis through expressions of anti-proliferative and pro-apoptotic genes (*e.g.*, *Tgfbi*, *Gadd45g*, *Dab*[236]). Furthermore, its activation inhibits the expression of several oncogenes, such as *c-Myc* and *HER-2*[237].

4.4.5 *Fascin*

Fascin, an actin-bundling protein, is a member of the cytoskeletal protein family which is highly expressed in neurons and dendritic cells[238]. Its normal functions are the maintenance of cell adhesion, motility and invasion. Many malignant epithelial cells have been shown to express elevated levels of fascin including those of liver, ovary, lung, pancreas and breast cancer[238-239] at levels which correlates with tumour invasiveness and metastasis[240]. Recent studies have demonstrated that STAT3 binds directly to fascin promoter sequences thus regulating its transcription[212a].

4.4.6 *NNMT*

Nicotinamide *N*-methyltransferase (NNMT) is a cytoplasmic enzyme that is thought to play a crucial role in the biotransformation of many drugs and xenobiotic compounds[241]. Recent studies have demonstrated aberrant expression of NNMT in several malignant tumours including thyroid cancer[242], glioblastoma[243], gastric cancer[244], renal carcinoma[245], lung cancer[246], pancreatic cancer[247], colorectal cancer[248] and ovarian carcinoma[249]. NNMT overexpression has been shown to correlate with cancer cell migration and tumour stage[250], and has been linked to cancer cell differentiation and poor prognosis[251]. An increasing body of evidence suggests that NNMT contributes to tumour development and progression through the induction of Complex I activity and ATP production, thereby supplying the tumour cells with energy[252]. Furthermore, studies have shown that NNMT promotes the activation of matrix metallopeptidase-2 *via* the PI3K/Akt signaling pathway, hence initiating tumour cell invasion and metastasis in renal clear cell carcinoma[253].

4.4.7 *VEGF*

VEGF is an essential vascular endothelial growth factor that stimulates vasculogenesis and angiogenesis in endothelial cells[254]. Besides its indispensability for vascular endothelial cells, VEGF is also important in bone formation[255], haematopoiesis[256] and wound healing[257]. Enhanced VEGF gene expression is induced by EGF (epithelial growth factor), TNF (tumour necrosis factor), IL-1 (interleukin-1) or hypoxia[258]. VEGF has been identified as a promoter of tumourigenesis and cancer progression in various human cancers[259]. The growth factor exerts this effect by stimulating tumour blood vessels growth to ensure its oxygen and nutrients supply[259].

4.4.8 *CREB5*

Cyclic adenosine monophosphate response element binding protein (CREB) is a proto-oncogenic TF[260] that plays a crucial role in cell survival[261]. Following external stimulation by hormones, growth factors and neuronal activity[262] CREB is phosphorylated by kinases, such as mitogen-activated protein kinase (MAPK) or protein kinase A (PKA)[262a, 263] resulting in activation of CREB dependent gene transcription. Several reports have demonstrated its overexpression and involvement in malignant diseases, including brain[264], breast[265], lung[266], prostate[267], bone marrow[268] and colorectal cancer[269].

4.4.9 *Elk-1*

Elk-1 belongs to the Ets (E twenty six) family of oncogenic TFs with critical roles in biological processes such as cell growth, differentiation and survival, haematopoiesis, angiogenesis, wound healing, cancer and inflammation[270]. It regulates transcription rather indirectly by interaction with co-regulatory proteins resulting in the combinatorial control of gene expression and enhanced specificity instead of directly binding to its DNA target sequences[271]. Elk-1 regulates AP-1 activation up-stream as it encodes genes for *c-Fos*, a component of the TF AP-1[272].

4.4.10 *AP-1*

The AP-1 (activating protein-1) dimeric TFs are composed of the subfamilies Jun, Fos, ATF, and MAF[273]. They execute and promote important biological functions including cell proliferation, migration, survival and apoptosis in specific cell types[59b]. AP-1 has been linked to cancer development and progression. Several reports have demonstrated aberrant activation of AP-1 in various malignant diseases such as melanoma[274], nasopharyngeal carcinoma[275], glioblastoma[276], liver cancer[277], breast cancer[278], sarcomas[279], Hodgkin lymphoma[60a] and myeloid leukaemia[280].

4.4.11 *p53*

P53, the most important tumour suppressor protein, is often referred to as 'guardian of the genome' due to its fundamental roles in maintaining genetic stability and prevention of tumour formation[281]. It executes defensive functions such as such as cell-cycle arrest, DNA synthesis and repair, apoptosis and energy metabolism[282] after DNA damage, nutrient starvation, heat shock, virus infection, hypoxia and oncogene activation[283]. Following its activation, p53 promotes induction of p21 which in turn forms a complex with Cdk2 leading to cell cycle arrest in phase G_1 in addition to the activation of GADD45 which stops cell division in G_2[284]. Disruptions in p53 expression have significant consequences to cells. On the one hand, overproduction of p53 is toxic to normal cells and on the other hand, a lack of p53 induces and promotes cancer formation and development[282].

4.5 Evaluation of SJG-136 in the breast cancer cell line MDA-MB-231

4.5.1 Results

4.5.1.1 SJG-136 down-regulated Bcl-2, cyclin D1, STAT1, STAT3, fascin and NNMT mRNA expression in PCR experiments

Polymerase Chain Reaction (PCR) analysis was used to evaluate changes in expression of *Bcl-2*, *cyclin D1*, *STAT3*, *STAT1*, *fascin* and *NNMT* in the breast cancer cell line MDA-MB-231. Gene expression profiles were compared before and after treatment with SJG-136 in the presence of bacterial lipopolysaccharide (LPS).

Initially, a general model was established in order to mimic the malignant disease. MDA-MB-231 cells were stimulated with 500 µg/mL LPS for 24 hours at 37 °C. LPS is the major component of the outer surface membrane of gram-negative bacteria which functions as an extremely effective stimulator of innate and natural immunity in eukaryotic cells[285]. LPS stimulates an extraordinary range of host responses (*i.e.*, shock, fever) by releasing cytokines including TNF and IL-1. Pre-treatment of MDA-MB-231 with LPS aimed to activate the transcription factor STAT3 resulting in elevation of STAT3-dependent gene expression. Following this, SJG-136 was added at 1 µM to the cells and incubated for another 24 hours at 37 °C. Finally, PCR was used to determine changes in gene expression profiles between stimulated-untreated and stimulated-treated cells.

PCR results are depicted in **Figure 4.3**. Activation of STAT3 and consequently, enhancement in STAT3-dependent gene expression, was successful using our established method. *Bcl-2*, *cyclin D1*, *STAT1*, *STAT3*, *fascin* and *NNMT* mRNAs were detected as single PCR products in MDA-MB-231 cells. For all analysed genes of interest an increase in band intensity was observed compared to the untreated cells (**Figure 4.3**, lane 1 *versus* lane 2) confirming STAT3 activation and STAT3-dependent gene over-expression. This result was surprising as it was expected to detect slightly elevated gene expression levels in the non-stimulated untreated cells (**Figure 4.3**, lane 1) as genes such as *Bcl-2* are known to be present at higher levels in human cancers[229a]. Following treatment of MDA-MB-231 with 1 µM SJG-136 for 24 hours at 37 °C decreased levels of mRNA expression

for all genes were detected (**Figure 4.3**, lane 2 *versus* lane 3) demonstrating down-regulation of STAT3 and STAT3-induced mRNA expression. The down-regulation was markedly pronounced for *Bcl-2*, *fascin* and *NNMT*. The consistent bands of the *GAPDH* control show equal loading of mRNA onto the agarose gel. These results suggest that SJG-136 may bind to consensus DNA sequences of the transcription factor STAT3 and, therefore, inhibit the expression of the STAT3-dependent genes *Bcl-2*, *cyclin D1*, *STAT3*, *fascin* and *NNMT*. Furthermore, the gel shown in **Figure 4.3** also demonstrates down-regulation of the transcription factor STAT1.

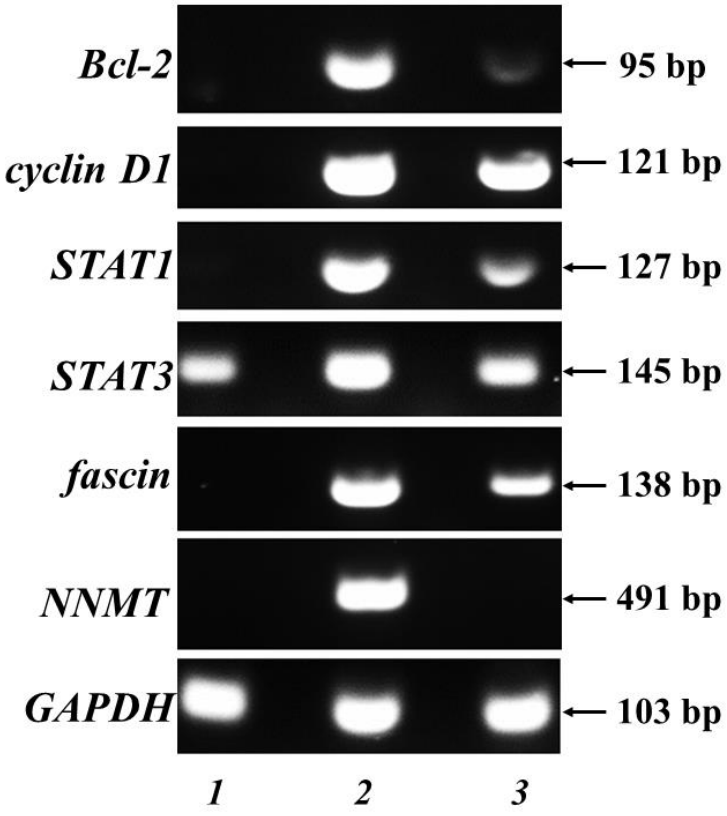


Figure 4.3: PCR detection of *Bcl-2*, *cyclin D1*, *STAT1*, *STAT3*, *fascin*, *NNMT* and control *GAPDH* expression in MDA-MB-231. Lane 1: untreated cells; lane 2: cells stimulated with 500 µg/mL LPS for 24 hours; lane 3: cells stimulated with 500 µg/mL LPS for 24 hours followed by 1 µM SJG-136 for 24 hours.

4.5.1.2 SJG-136 showed concentration dependent effects on Bcl-2, cyclin D1, STAT1, STAT3, fascin, and NNMT mRNA down-regulation

To investigate the concentration-dependent activity of SJG-136 on the expression of STAT3-dependent genes and STAT1 in MDA-MB-231 cells, the ligand was tested at various concentrations (0.001 - 1 µM) following the protocol described in Section 4.5.1.1. The results shown in **Figure 4.4** demonstrate down-regulation of *Bcl-2*, *STAT1*, *STAT3* and *NNMT* at 1 µM and 0.1 µM but no visible down-regulation at concentrations of 0.01 µM and 0.001 µM. C*yclin D1* and *fascin* were slightly down-regulated at 1 µM but no detectable down-regulation was achieved at 0.1 µM, 0.01 µM and 0.001 µM SJG-136. From these data we can propose that the effect of SJG-136 on STAT3 consensus sequences and STAT3-dependent gene expression is concentration-dependent.

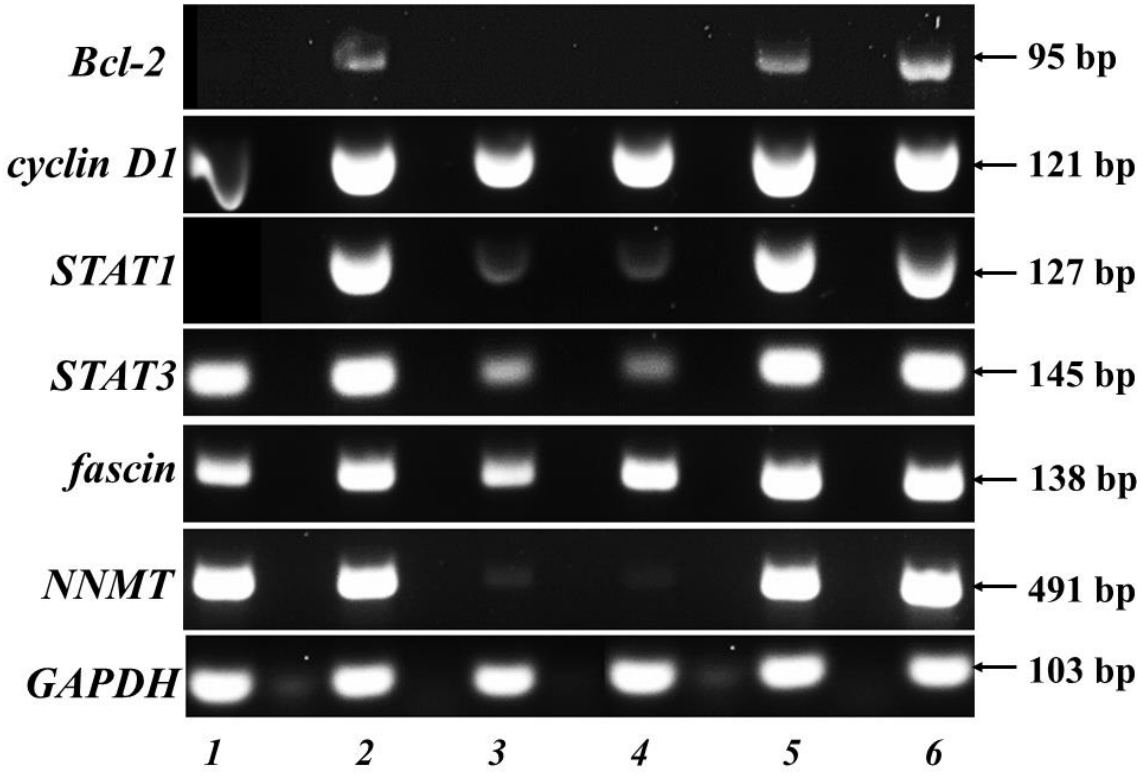


Figure 4.4: PCR detection of *Bcl-2*, *cyclin D1*, *STAT1*, *STAT3*, *fascin*, *NNMT* and control *GAPDH* expression in MDA-MB-231. Lane 1: untreated cells; lane 2: cells stimulated with 500 µg/mL LPS for 24 hours; lane 3: cells stimulated with 500 µg/mL LPS for 24 hours followed by 1 µM SJG-136 for 24 hours; lane 4: cells stimulated with 500 µg/mL LPS for 24 hours followed by 0.1 µM SJG-136 for 24 hours; lane 5: cells stimulated with 500 µg/mL LPS for 24 hours

followed by 0.01 µM SJG-136 for 24 hours; lane 6: cells stimulated with 500 µg/mL LPS for 24 hours followed by 0.001 µM SJG-136 for 24 hours.

4.5.1.3 SJG-136 significantly down-regulated Bcl-2, VEGF, p53, STAT3, CREB5, and Elk-1 mRNA expression in qPCR study

Quantitative PCR (qPCR) was used to verify the observed changes in our PCR study in gene expression profiles after the cells were stimulated with LPS and treated with SJG-136. Nine genes, *cyclin D1*, *VEGF*, *AP-1*, *Bcl-2*, *p53*, *CREB5*, *STAT3*, *survivin* and *Elk-1* were chosen for evaluation by qPCR and normalised against the reference gene *β-actin*. Statistical analysis was performed using 1-way ANOVA followed by Turkey's Multiple Comparison post-test. *P*-values < 0.05 were considered as significant. qPCR results are shown in **Figure 4.5**. The qPCR data showed that a total of six genes (*VEGF*, *Bcl-2*, *p53*, *CREB5*, *STAT3* and *Elk-1*) were significantly down-regulated with fold changes ranging from 2.3- to 2.0-fold after the cells were stimulated with LPS and treated with SJG-136 (**Figure 4.5**, blue column) compared to the LPS-stimulated and non-treated cells (**Figure 4.5**, green column).

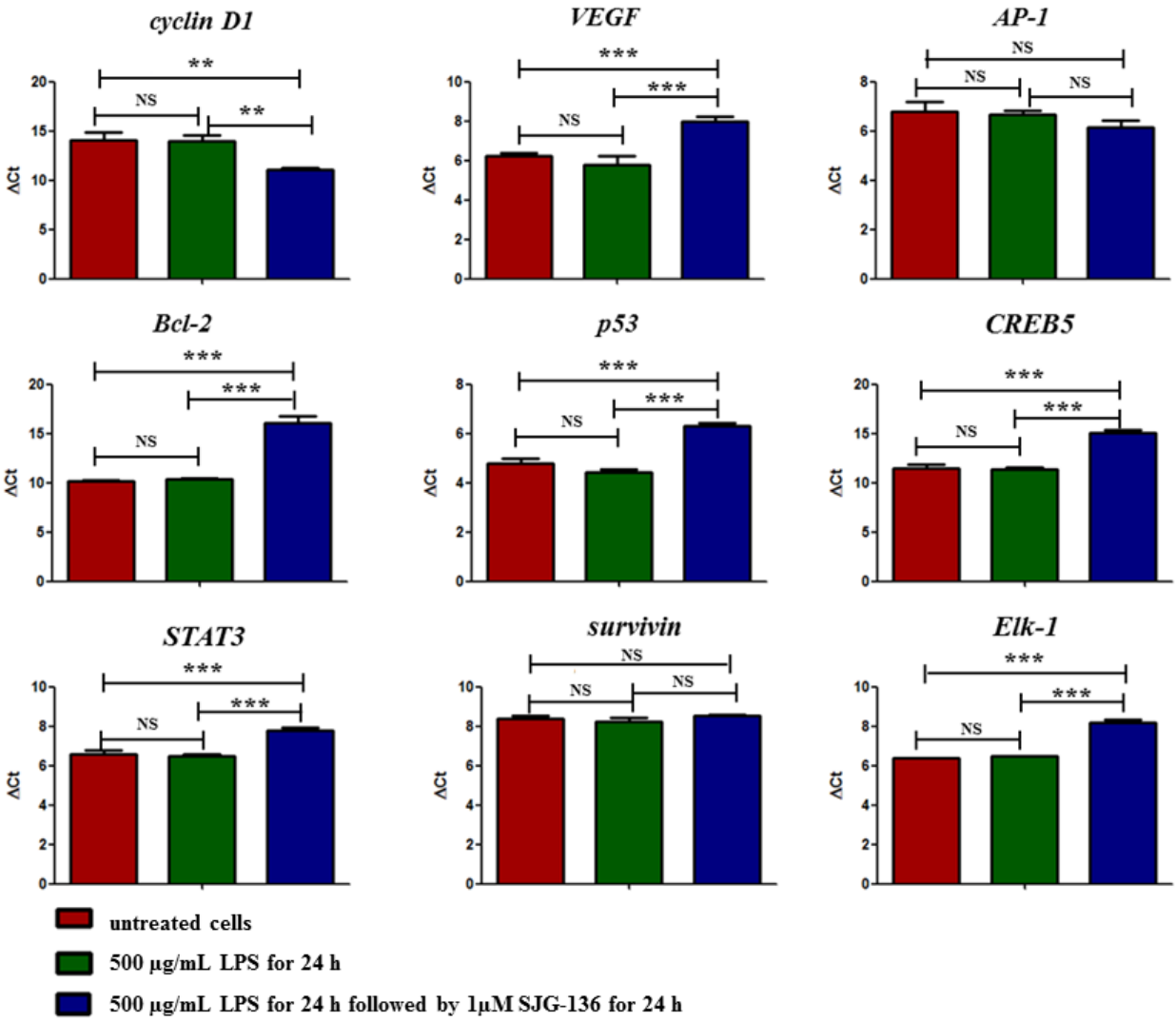


Figure 4.5: The effect of SJG-136 on the expression of *cyclin D1*, *VEGF*, *AP-1*, *Bcl-2*, *p53*, *CREB5*, *STAT3*, *survivin* and *Elk-1* mRNA in MDA-MB-231. All data are mean ± SD. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, NS = not significant. Statistical analysis was performed using a 1-way ANOVA followed by a Turkey's Multiple Comparison post-test, n = 3. An increase in ΔCt equates to a decrease in gene expression.

Analysis of variance with a *p*-value of 0.15 for *AP-1* and a *p*-value of 0.11 for *survivin* were found not to be significant (**Figure 4.5**). Surprisingly, the most striking change in gene expression before and after ligand treatment was observed for *cyclin D1*, a major

regulator of the cell cycle (**Figure 4.5**). A possible explanation for this surprising observation shall be discussed in detail in Section 4.7.

4.5.1.4 No significant over-expression of STAT3-dependent genes was achieved after stimulation with LPS

No significant up-regulation of STAT3-dependent gene expression was achieved after stimulation of MDA-MB-231 with LPS (red *versus* green column, **Figure 4.5**). This was not consistent with the results obtained from the PCR study, in which for each case, over-expression of STAT3 was observed (lane 1 *versus* lane 2, **Figure 4.3**). Possible reasons for this observation will be discussed in Section 4.7 in detail.

4.5.1.5 SJG-136 inhibited cyclin D1 protein expression and up-regulated phosphorylated STAT1 protein expression

To further validate the PCR findings, Western blotting was performed to correlate changes in gene and protein expression levels in LPS-stimulated non-treated and LPS-stimulated SJG-136-treated MDA-MB-231 cells. Cyclin D1 and Bcl-2 were chosen for the Western blotting study due to their crucial involvement in tumour development and progression[228, 233a]. STAT1 was selected for the study because of its tumour suppressive features and to investigate the selectivity of SJG-136 toward STAT3. Statistical analysis was performed using 1-way ANOVA followed by Turkey's Multiple Comparison post-test. *P*-values < 0.05 were considered as significant. Western blot results are shown in **Figure 4.6**, **4.7** and **4.8**. All proteins were detected as single bands of the following sizes: cyclin D1 ~36 kDa[286], Bcl-2 ~28 kDa[287], phosphorylated STAT1 (pSTAT1) ~87 kDa[236b], total STAT1 (tSTAT1) ~87 kDa[288], phosphorylated STAT3 (pSTAT3) ~88 kDa[236b], total STAT3 (tSTAT3) ~88 kDa[288] and α-tubulin ~50 kDa[289]. In all experiments, no non-specific binding of the antibodies (appearance of multiple bands of different sizes) was detected on the membranes, confirming specific binding of the primary antibodies to the proteins of interest. **Figure 4.6** shows visible down-regulation of cyclin D1 and up-regulation of pSTAT1 protein expression after the cells were stimulated with LPS and treated with SJG-136 (lane 2 *versus* lane 3). The subsequent statistical analysis confirmed significant down-regulation in cyclin D1 protein expression in LPS-stimulated and SJG-136-treated cells (**Figure 4.7**, green *versus* blue column). Accordingly, the

pSTA1:tSTAT1 ratio significantly increased after the cells were stimulated with LPS and treated with SJG-136 (**Figure 4.8**, green *versus* blue column).

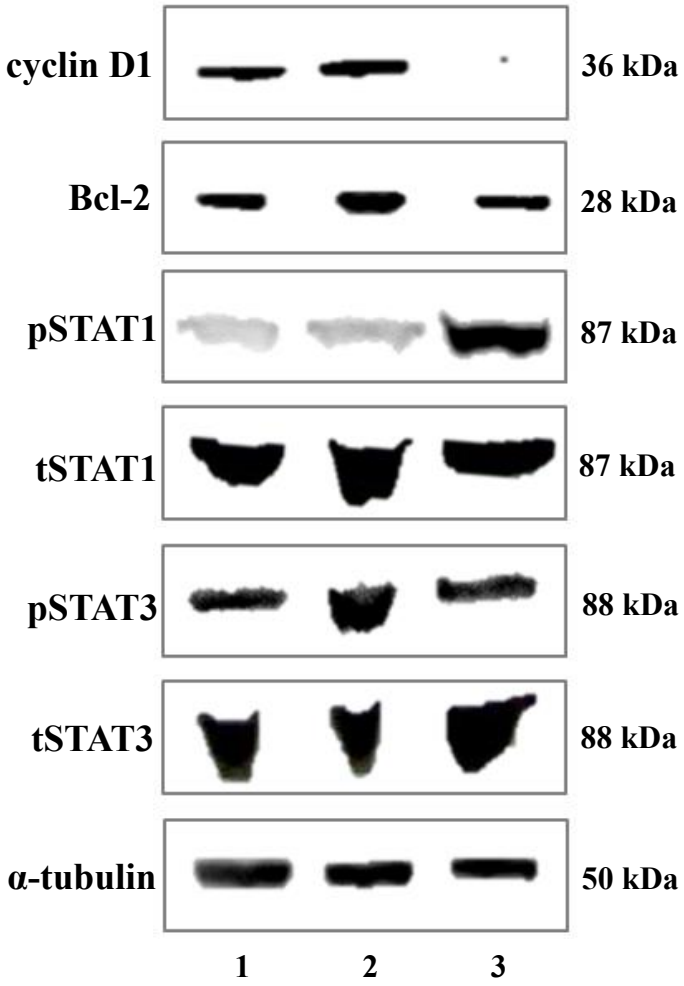


Figure 4.6: Western blots of MDA-MB-231 for cyclin D1, Bcl-2, pSTAT1, tSTAT1, pSTAT3, tSTAT3 and α-tubulin expression. Molecular weights: cyclin D1 ~36 kDa, Bcl-2 ~28 kDa, pSTAT1 ~87 kDa, tSTAT1 ~87 kDa, pSTAT3 ~88 kDa, tSTAT3 ~88 kDa, α-tubulin ~50 kDa. Lane 1: untreated cells; lane 2: 500 µg/mL LPS for 24 hours; lane 3: 500 µg/mL LPS for 24 hours followed by 1 µM SJG-136 for 24 hours.

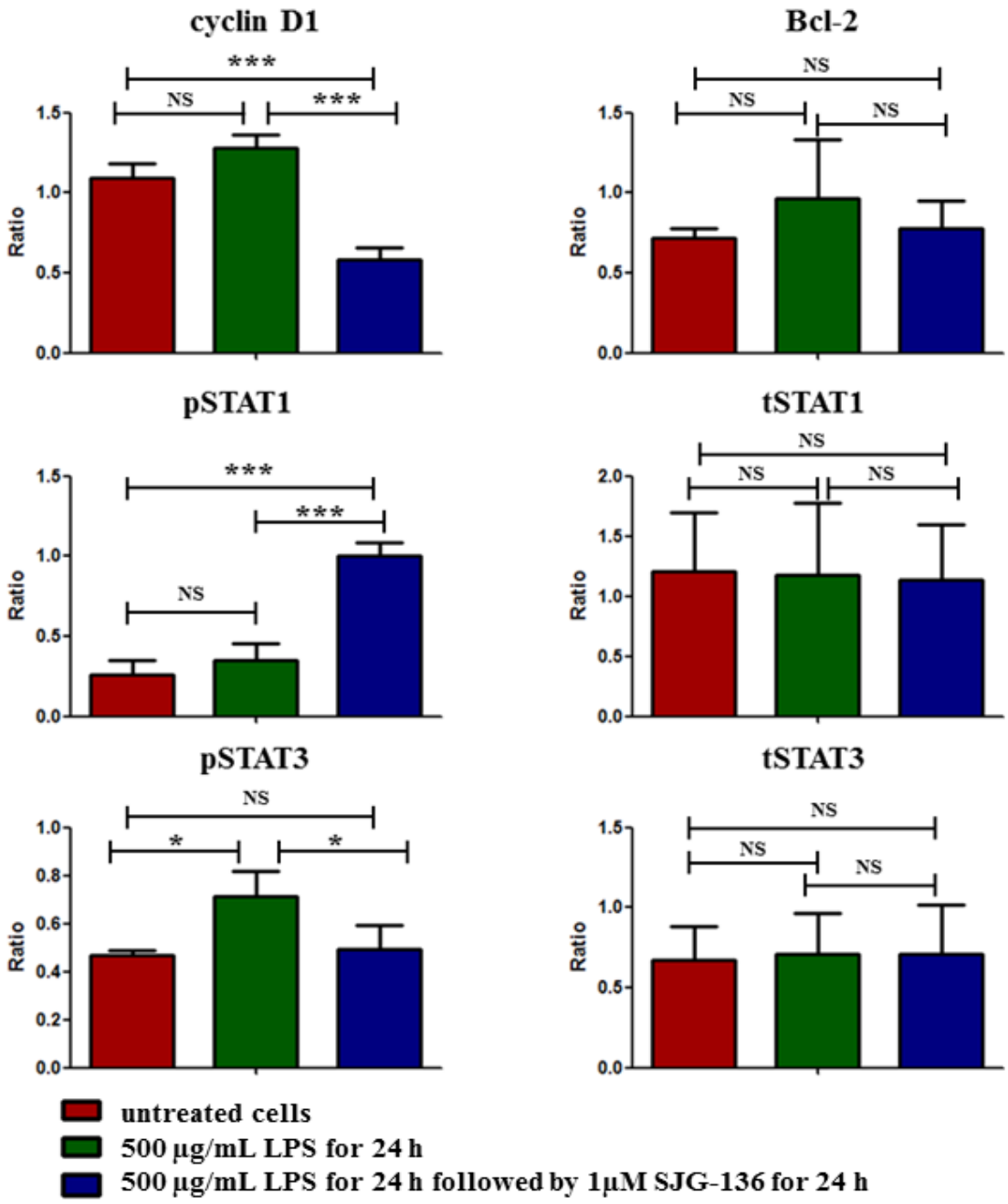


Figure 4.7: The effect of SJG-136 on the protein expression of cyclin D1, Bcl-2, pSTAT1, tSTAT1, pSTAT3 and tSTAT3 in MDA-MB-231. All data are mean ± SD. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, NS = not significant. Statistical analysis was performed using a 1-way ANOVA followed by a Turkey's Multiple Comparison post-test, n = 3.

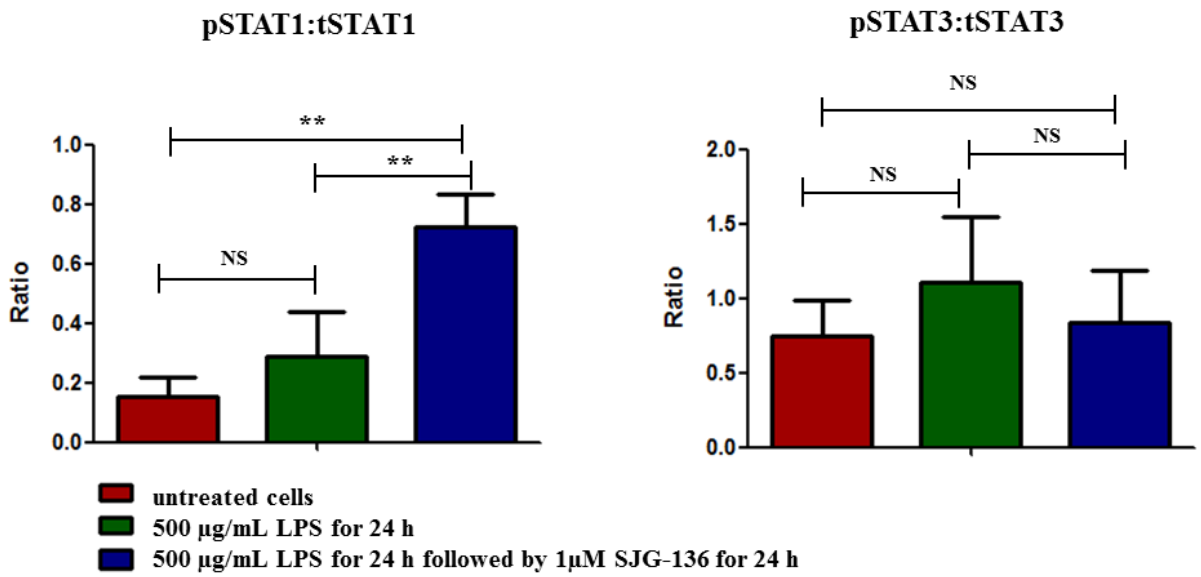


Figure 4.8: The effect of SJG-136 on the protein expression of pSTAT1, tSTAT1, pSTAT3 and tSTAT3 in MDA-MB-231. All data are mean ± SD. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, NS = not significant. Statistical analysis was performed using a 1-way ANOVA followed by a Turkey's Multiple Comparison post-test, n = 3. STAT1 phosphorylation was calculated by dividing the pSTAT1 intensity by the tSTAT1 intensity to produce the pSTAT1:tSTAT1 ratio. The same procedure was applied for the pSTAT3:tSTAT3 ratio.

4.5.1.6 SJG-136 did not change levels of Bcl-2, phosphorylated STAT3 and total STAT1 and STAT3 proteins

No visible down- or up-regulation was detected for Bcl-2, tSTAT1, pSTAT3 and tSTAT3 (**Figure 4.6**) using the LPS model. Additionally, statistical analysis confirmed no significant changes in Bcl-2 and tSTAT1 protein expression (**Figure 4.7**, green *versus* blue column). The pSTAT3:tSTAT3 ratio did not increase after the cells were stimulated with LPS and treated with SJG-136 confirming no significant changes in phosphorylated STAT3 protein expression (**Figure 4.8**, green *versus* blue column). In agreement with the qPCR results, no over-expression of any protein investigated was achieved after stimulation with LPS, suggesting no increased expression of STAT3-dependent genes (see Discussion in Section 4.7) (**Figure 4.7** and **Figure 4.8**, red *versus* green column).

4.6 Evaluation of SJG-136 in the colon carcinoma cell line HT-29

4.6.1 Results

4.6.1.1 SJG-136 down-regulated VEGF, cyclin D1, CREB5, AP-1 and p53 mRNA expression in PCR experiments

PCR analysis was undertaken to investigate differences in expression of AP-1 down-stream target genes *VEGF*, *cyclin D1*, *CREB5*, *Elk-1*, *AP-1* and *p53* in the colon carcinoma cell line HT-29 before and after incubation with SJG-136 to correlate the observations made during the biophysical analysis of SJG-136. Using the developed model, 30 ng/mL tumour necrosis factor α (TNF-α) was added to HT-29 cells and incubated for 24 hours at 37 °C with the aim to increase the expression of AP-1. TNF-α is a proinflammatory cytokine with a variety of biological functions including cell proliferation, differentiation, apoptosis, lipid metabolism, and coagulation[290]. TNF-α has been demonstrated to induce the AP-1 signalling pathway by activation of specific kinases required for phosphorylation c-Jun and ATF2 which are part of AP-1[291]. Following this, cells were treated with 1 µM SJG-136 for another 24 hours at 37 °C in the presence of TNF-α and gene expression profiles were compared using PCR analysis. The obtained results are shown in **Figure 4.9**. Over-expression of AP-1 and consequently over-expression of AP-1-dependent genes, was successful after the cells were stimulated with TNF-α (**Figure 4.9**, lane 1 *versus* lane 2). *VEGF*, *cyclin D1*, *CREB5*, *Elk-1*, *AP-1* and *p53* mRNAs were detected as single PCR products in HT-29 cells. **Figure 4.9** shows an increase in band intensity compared to untreated cells for all analysed genes (lane 1 *versus* lane 2) confirming activation of AP-1 and over-expression AP-1-dependent genes. After treatment of HT-29 with TNF-α and 1 µM SJG-136 for 24 hours at 37 °C the mRNA expression levels decreased for all genes as visible on the decrease of band intensities (**Figure 4.9**, lane 2 *versus* lane 3). This result demonstrates down-regulation of AP-1 and AP-1-depentent gene expression. The consistent bands of the control *GAPDH* confirm equal loading of mRNA onto the agarose gel. Interestingly, no visible down-regulation was observed for *Elk-1* as shown in **Figure 4.9**, lanes 2 & 3. The band intensities of the cells treated with only TNF-α and the cells treated with TNF-α and SJG-136 are similar. A possible explanation for this observation will be discussed later. These PCR results

indicate that SJG-136 may be able to bind to the DNA consensus sequences of AP-1, thereby inhibiting the expression of AP-1-dependent genes such as *VEGF*, *cyclin D1*, *CREB5*, *Elk-1*, *AP-1* and *p53*.

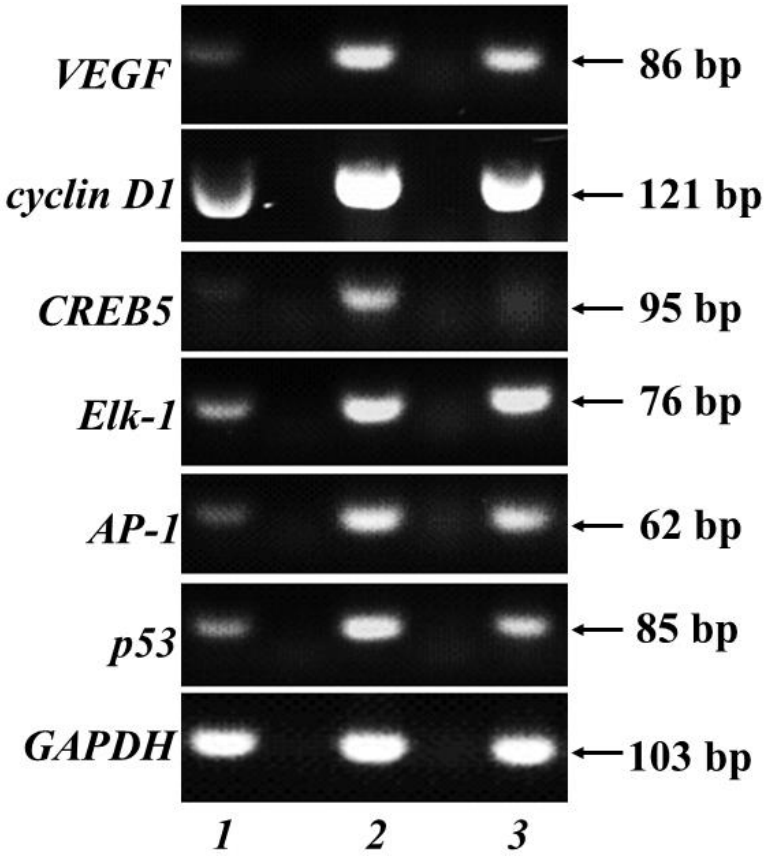


Figure 4.9: PCR detection of *VEGF*, *cyclin D1*, *CREB5*, *Elk-1*, *AP-1*, *p53* and control *GAPDH* expression in HT-29. Lane 1: untreated cells; lane 2: cells stimulated with 30 ng/mL TNF-α for 24 hours; lane 3: cells stimulated with 30 ng/mL TNF-α for 24 hours followed by 1 μM SJG-136 for 24 hours.

4.6.1.2 SJG-136 significantly down-regulated Bcl-2, p53, CREB5 and survivin mRNA expression in qPCR study

Validation of the HT-29 PCR results was carried out using qPCR. Differences in mRNA expression levels of *cyclin D1*, *VEGF*, *AP-1*, *Bcl-2*, *p53*, *CREB5*, *STAT3*, *survivin* and *Elk-1* were quantified in TNF-α-stimulated non-treated cells and TNF-α-stimulated and SJG-136-treated cells. Results were normalised against the reference gene *β-actin*. Statistics was conducted using 1-way ANOVA followed by Turkey's Multiple

Comparison post-test. *P*-values < 0.05 were considered as being significant. Obtained qPCR results are shown in **Figure 4.10**. Statistical analysis showed that the expression of *Bcl-2*, *p53*, *CREB5* and *survivin* were significantly down-regulated after TNF-α-stimulated cells were treated with SJG-136 (**Figure 4.10**, green *versus* blue column). The fold changes in mRNA expression levels decreased from 3.7- to 2.9 (TNF-α-stimulated non-treated cells) to 2.6- to 2.0-fold (TNF-α-stimulated and SJG-136-treated cells) (**Figure 4.10**, green *versus* blue column).

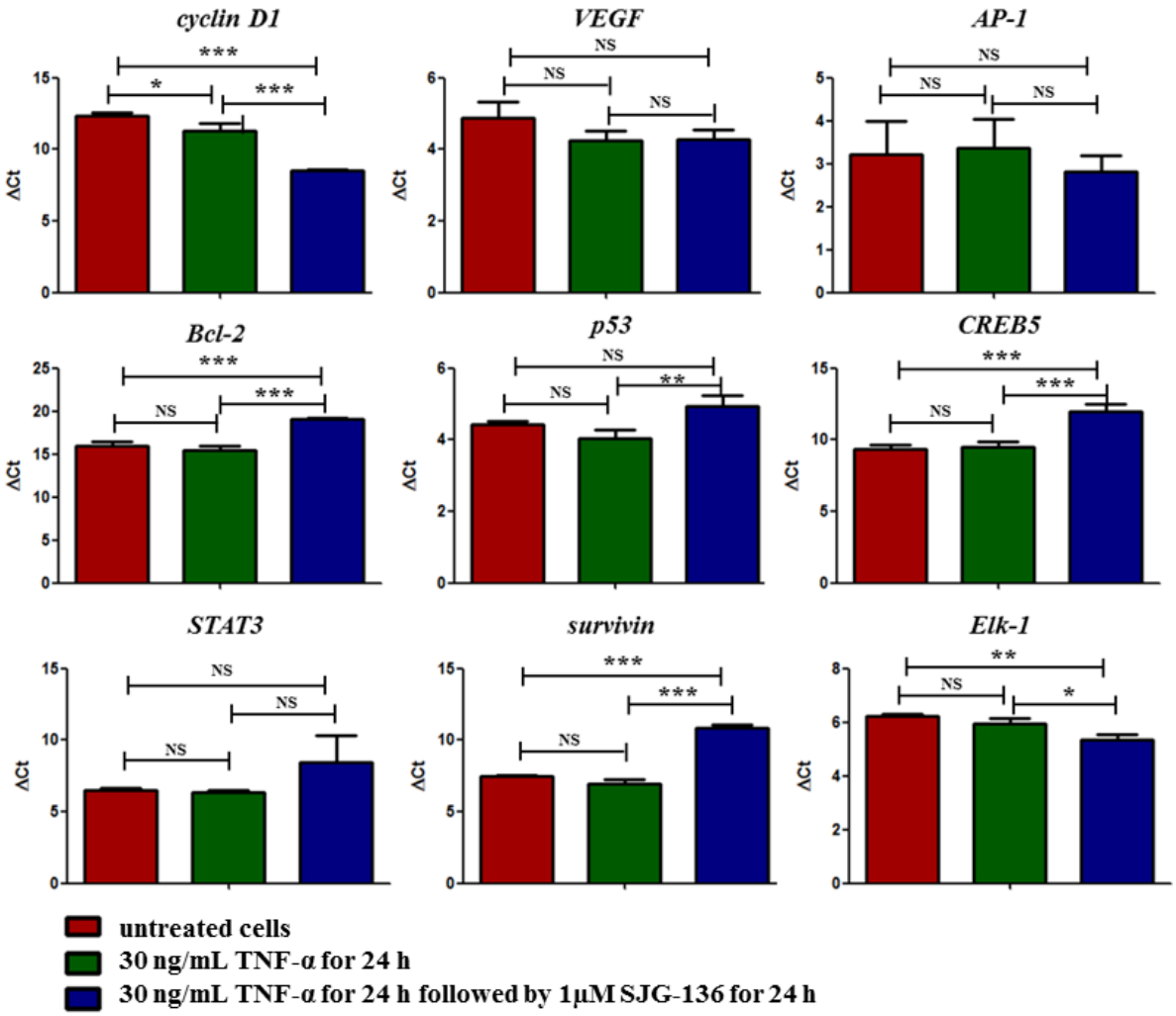


Figure 4.10: The effect of SJG-136 on the expression of *cyclin D1*, *VEGF*, *AP-1*, *Bcl-2*, *p53*, *CREB5*, *STAT3*, *survivin* and *Elk-1* mRNA in HT-29. All data are mean ± SD. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, NS = not significant. Statistical analysis was performed using a 1-way ANOVA followed by a Turkey's Multiple Comparison post-test, n = 3. An increase in ΔCt equates to a decrease in expression.

Statistical analysis using 1-way ANOVA showed no significant differences in gene expression for *VEGF*, *AP-1* and *STAT3* (**Figure 4.10**, green *versus* blue column).

Cyclin D1 produced a similar change in expression in HT-29 cells as that observed in MDA-MB-231 with an increased expression after stimulation with TNF-α and treatment SJG-136 (**Figure 4.10**, green *versus* blue column). Furthermore, *Elk-1* was also found to be significantly up-regulated after treatment with SJG-136 with an increase in expression from 3.4- to 4.4-fold (**Figure 4.10**, green *versus* blue column).

4.6.1.3 No significant over-expression of AP-1-dependent genes was achieved after stimulation with TNF-α

Lastly, no significant over-expression of AP-1-dependent genes was achieved after stimulation with TNF-α (**Figure 4.10**, red *versus* green column). However, *cyclin D1* showed significant up-regulation in its expression after stimulation of HT-29 with TNF-α. This will be discussed in Section 4.7 in detail.

4.6.1.4 SJG-136 significantly inhibited cyclin D1 protein synthesis in Western blot experiments

Protein expression levels of cyclin D1, p53 and VEGF in HT-29 were determined using Western blotting to investigate whether any correlation is present between the gene expression and protein expression levels after treatment with SJG-136. HT-29 cells were stimulated with TNF-α for 24 hours and treated for another 24 hours with SJG-136 in the presence of TNF-α. Analysis of variance was carried out using 1-way ANOVA followed by Turkey's Multiple Comparison post-test. *P*-values < 0.05 were taken as being significant. Western blot results are depicted in **Figure 4.11** and **Figure 4.12**). Analysed proteins were detected as single bands with the following sizes: cyclin D1 ~36 kDa[286], p53 ~53 kDa[292], VEGF ~43 kDa[293] and α-tubulin (loading control) ~50 kDa[289]. No additional protein bands were present on the membranes, thus demonstrating a specific interaction of the primary antibodies with the investigated proteins. As visible in **Figure 4.11**, only cyclin D1 was significantly down-regulated in its expression after the cells were stimulated with TNF-α and treated with SJG-136.

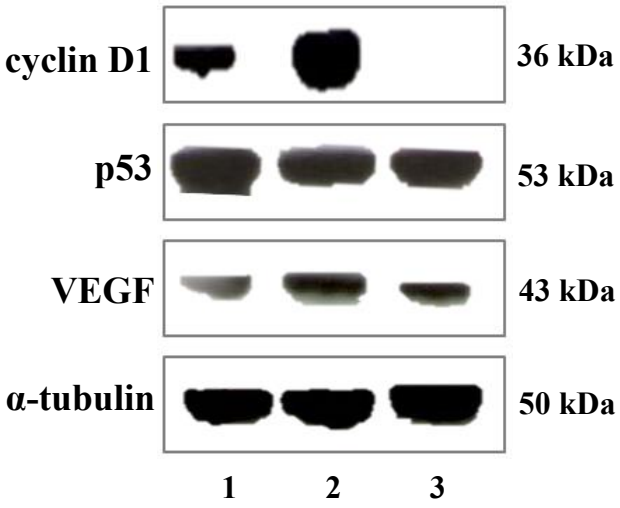


Figure 4.11: Western blot analysis of cyclin D1, p53 and VEGF protein expression in HT-29, using α-tubulin as a loading control (n = 3). Detected kDa: cyclin D1 ~36 kDa, p53 ~53 kDa, VEGF ~43 kDa, and control α-tubulin ~50 kDa. Lane 1: untreated cells; lane 2: 30 ng/mL TNF-α for 24 hours; lane 3: 30 ng/mL TNF-α for 24 hours followed by 1 µM SJG-136 for 24 hours.

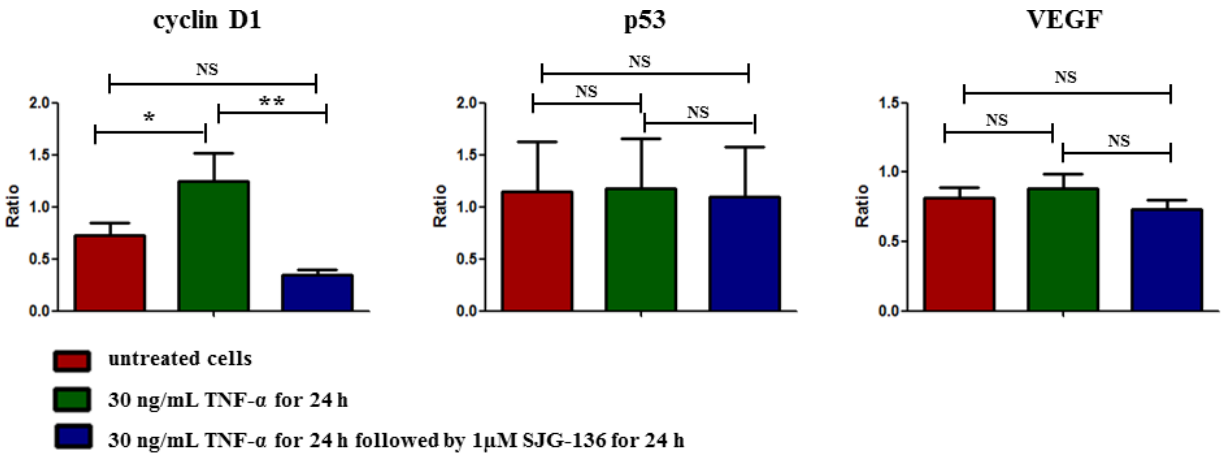


Figure 4.12: The effect of SJG-136 on the protein expression of cyclin D1, p53 and VEGF in HT-29. All data are mean ± SD. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, NS = not significant. Statistical analysis was performed using a 1-way ANOVA followed by a Turkey's Multiple Comparison post-test, n = 3.

4.6.1.5 SJG-136 had no effect on VEGF and p53 protein expression

No significant down-regulation was observed in p53 and VEGF protein expression after stimulating of HT-29 cells with TNF-α followed by treatment with SJG-136 (**Figure 4.12**, green *versus* blue column). These results were confirmed by subsequent statistical analysis using 1-way ANOVA followed by Turkey's Multiple Comparison post-test which demonstrated a significant decrease in cyclin D1 protein expression (p = 0.0185) and no significant changes in p53 (p = 0.0726) and VEGF (p = 0.1688) protein expression (**Figure 4.12**). In agreement with the qPCR data, no protein over-expression was observed for p53 and VEGF after the cells were treated with TNF-α. This result suggests that the activity of AP-1 may have been induced but that the consequences upon p53 and VEGF protein expression have not occurred after 24 hours (**Figure 4.12**, red *versus* green column). Time-course experiments could help to see any effect on protein expression after treatment with SJG-136 which is not apparent at the single time-point of 24 hours used during the applied TNF- α model.

4.7 Discussion

The biological evaluation of SJG-136 on AP-1- and STAT3-dependent gene expression on the human colon carcinoma cell line HT-29 and the breast cancer cell line MDA-MB-231 have shown that SJG-136 may bind to the response elements on DNA of AP-1 and STAT3, thereby inhibiting AP-1- and STAT3 dependent gene expression. The PCR study demonstrated down-regulation of *cyclin D1*, VEGF, *CREB5*, *p53* and *AP-1* mRNA expression levels after HT-29 cells were stimulated with TNF-α and treated with SJG-136. Similarly, *Bcl-2*, *cyclin D1*, *STAT3*, *fascin* and *NNMT* were found to be down-regulated in their expression after the MDA-MB-231 cells were stimulated with LPS and treated with SJG-136 during the same PCR experiment. The qPCR study showed significant down-regulation of *Bcl-2*, *p53*, *CREB5* and *survivin* mRNA expression after stimulation of HT-29 with TNF-α and treatment with SJG-136 and a significant decrease mRNA expression of *VEGF*, *Bcl-2*, *p53*, *CREB5*, *STAT3* and *Elk-1* in LPS-stimulated and SJG-136-treated MDA-MB-231. Finally, Western blotting results demonstrated a significant decrease in cyclin D1 protein expression in both, TNF-α-stimulated and SJG-136-treated HT-29 and LPS-stimulated and SJG-136-treated MDA-MB-231 cells.

It is clear however, that the PCR results where SJG-136 appeared to down-regulate the majority of the analysed genes are not fully consistent with Western blot data, where only down-regulation of cyclin D1 protein expression was observed in both the HT-29 and MDA-MB-231 cells. A possible explanation for this inconsistency may be the fact that changes in mRNA expression do not always correlate with changes in protein expression. The process of mRNA expression is very rapid and transient. mRNA turnover happens within minutes and does not last for longer than a few hours[214a]. In contrast, the process of protein expression is very slow. Some proteins have half-lives of ~16 hours such as Bcl-2 ($t_{1/2}$ ~22 hours[294]) and STAT1 ($t_{1/2}$ ~18 hours[295]), whereas the average is 4 – 8 hours. This means that a number of half-lives are required in order to observe a reduction in protein expression. Moreover, a down-regulation in mRNA expression should lead to a decrease in protein expression. However, this process can take a while. As such, any changes in mRNA expression may not translate to changes in protein expression over the course of the experiment. Therefore, the single time-point of 24 hours used in the LPS model may not be long enough to detect any correlation between the mRNA and protein

expression levels. Western blot time-course studies adjusted to the half-lives of the proteins of interest, may help to observe any effect on protein expression after treatment of the cells with SJG-136

Furthermore, post-translational regulation mechanisms may contribute to the observed differences in mRNA and protein expression levels during the study. Additional modifications made on the proteins can be phosphorylation, acetylation, ubiquination, methylation, glycosylation, and nitrosylation which all at the end determine whether or not the protein will be active[214b].

In addition, the PCR and qPCR data did not correlate after the cells were stimulated with TNF-α (HT-29) and LPS (MDA-MB-231). This step aimed to induce an increased expression of AP-1- and STAT3-dependent genes. However, during the PCR study over-expression was achieved for all studied genes in both HT-29 and MDA-MB-231, visible on the increased band intensity, whereas the qPCR study demonstrated significant over-expression only of *cyclin D1* in HT-29 after the cells were treated with TNF-α. Moreover, the PCR results showed a visible down-regulation of *cyclin D1* after treatment with SJG-136 in both cell lines whereas the qPCR data demonstrated significant over-expression of *cyclin D1* after the cell were incubated with the ligand.

This discrepancy is likely due to some major limitations of the PCR. PCR is semi-quantitative, as the used DNA intercalators such as ethidium bromide lack specificity. They bind to any double-stranded DNA which may result in misleading results. Furthermore, the staining with intercalators can prevent further amplification, thereby complicating quantification of the genes of interest. Moreover, PCR is non-linear and may not be exponential for the first several cycles, and the reaction will plateau at a certain point because essential reagents such as dNTP, enzyme and primers get depleted. This results in saturation of the UV or colourimetric signal which will not increase in intensity when the reaction proceeds further. This all affects subsequent quantitative analysis. Therefore, end-point PCR is not a very useful tool to accurately determine whether a particular gene is expressed or not in a given sample. In order to precisely quantify any changes in gene expression, the amount of product formed needs to be visualised during the log phase of the reaction which is the case in qPCR.

The observations made for phosphorylated STAT1 during Western blot analysis are interesting. The results showed a significant over-expression of phosphorylated STAT1 protein after the MDA-MB-231 cells were treated with LPS and SJG-136 suggesting its increased expression in tumour cells. This enhanced expression of the active form of STAT1 protein may reflect the tumour suppressor ability of STAT1[296].

4.8 Conclusion

In conclusion, the biological evaluation of the effect of SJG-136 on AP-1- and STAT3-dependent gene expression has demonstrated that the ligand may be able to interact with the AP-1 and STAT3 cognate sequences and thereby inhibit the expression of AP-1- and STAT3-dependent genes involved in tumourigenesis and progression, such as *Bcl-2*, *fascin*, *NNMT*, *VEGF*, *CREB5* and *p53*. This observed down-regulation may be due to interaction with the consensus sequences of the transcription factors AP-1 and STAT3 which was observed during the biophysical study.

The Western blot data have not clearly confirmed the ability of SJG-136 to down-regulate AP-1- and STAT3-dependent protein expression as, from the panel of investigated proteins, only cyclin D1 was found to be significantly down-regulated by SJG-136. Possible reasons for the observed discrepancies between the PCR and Western blot data are discussed in Section 4.7 in detail. This will be investigated in future studies in detail as outlined in Chapter 6.

Chapter 5: DNA-binding site analysis of SJG-136 within the AP-1 consensus sequence

5.1 Background

During a study of the interaction of SJG-136 (**2**, **Figure 5.1**) with various DNA transcription factor recognition sequences[191b] (see also Chapter 3), it was observed that SJG-136 (**2**) rapidly formed an adduct with the AP-1 hairpin DNA sequence (**Figure 5.2**), with a major new peak appearing in the chromatogram at retention time (RT) 7.53 min after 24 hours of incubation, and with reaction complete by 24 hours (**Figure 3.5A** and **3.5B**, see also Section 3.2.1). The stoichiometry of the adduct was confirmed as 1:1 SJG-136/AP-1 by MALDI-TOF-MS with an observed mass of 6351.2 m/z (theoretical mass: 6350.41 m/z) (**Figure 3.5C**, Section 3.2.1). Given that it has previously been shown[96] that **2** can form mono-adducts and intra- and interstrand cross-links of different lengths according to the following rank order of preference: Pu-GAATG-Py > Pu-GATC-Py >> Pu- GATG-Py and Pu-GAATC-Py[205], it was initially assumed that the adduct formed was most likely to be either the extended G7-G17 3'-GTAAC-5' interstrand cross-link, a G7 or G17 mono-alkylated adduct, or a combination of one or more of these (**Figure 5.2**).

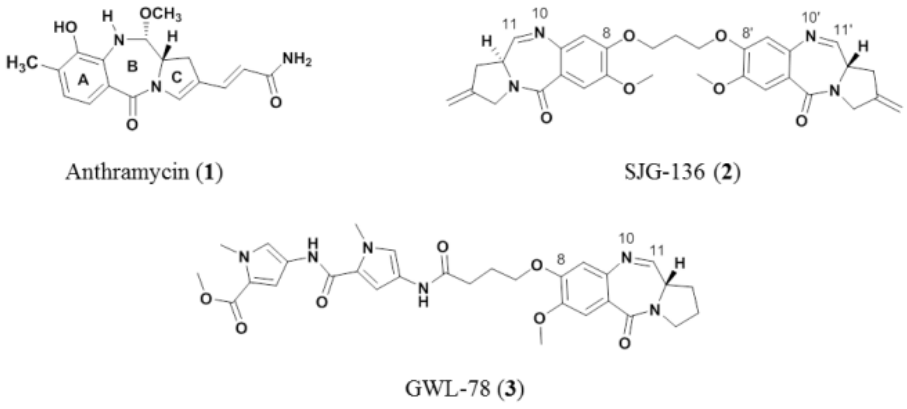


Figure 5.1: Structures of the naturally occurring PBD anthramycin (**1**), the synthetic PBD Dimer SJG-136 (**2**), and the synthetic C8-*bis*-pyrrole PBD Conjugate GWL-78 (**3**).

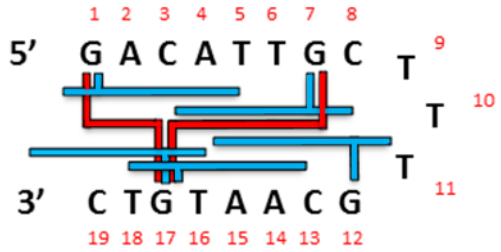


Figure 5.2: Schematic diagram of the possible interstrand (red) and mono-alkylated (blue) adducts that could form within the parent AP-1 hairpin sequence.

5.2 Objectives

The objective of this study was to investigate what particular type of adduct SJG-136 had formed during the HPLC study with the AP-1 sequence using the RP-HPLC/MALDI-TOF method developed in Chapter 2. In order to examine this, a range of inosine modified DNA hairpins und duplexes were designed (shown in the individual sections) which are based on the same AP-1 sequence but allow SJG-136 to form mono-alkylated and cross-linked adducts with the AP-1 sequence. The guanines were consecutively replaced with inosine bases to remove the nucleophilic C2-NH_2 functionalities and prevent covalent interaction of SJG-136.

5.3 Interaction of SJG-136 with inosine modified AP-1 hairpin sequences to study cross-link formation

The substitution patterns of the hairpins AP-1 1, 17I, AP-1 7, 17I, AP-1 7, 12I and AP-1 1, 12I (**Figure 5.3**) allowed the possibility of cross-linking at G7 and G12, G1 and G12, G1 and G17, and G7 and G17 due to the replacement of two out of the four guanines with non-nucleophilic inosines, thereby allowing the possibility of interstrand cross-link formation and mono-alkylation by SJG-136.

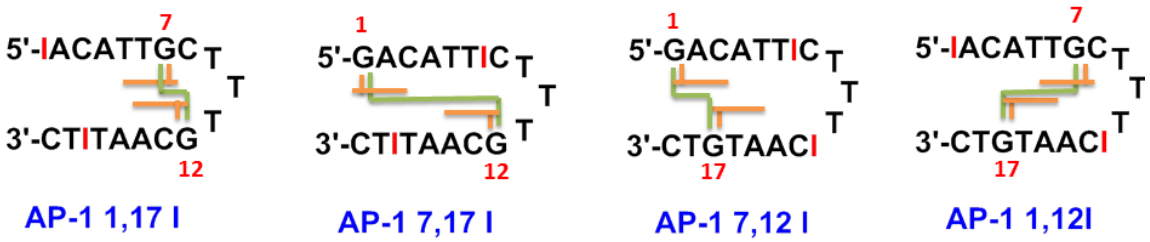


Figure 5.3: Structures of inosine modified AP-1 hairpins used to study cross-link formation of SJG-136 which are based on the same AP-1 sequence but with two of the four guanines replaced with inosines in each case to study cross-linking.

5.3.1 Results and discussion

Working solutions of inosine modified hairpins were prepared according to the protocol described in Section 7.2.1.4. SJG-136 working solutions were prepared as described in Section 7.2.1.9. SJG-136/DNA complexes were prepared according to the protocol described in Section 7.2.1.10. Briefly, a SJG-136 working solution of 100 µM was added to a DNA working solution of 25 µM in a 4:1 ratio (SJG-136/DNA). The mixture was agitated and incubated for 24 hours at 25 °C. Following the incubation, the samples were subjected to ion pair RP-HPLC/MALDI-TOF-MS analysis as described in Sections 7.2.2 and 7.2.4 in detail.

HPLC analysis of annealed AP-1 1, 17I (**Figure 5.3**), with guanines G7 and G12 available showed a single peak at retention time (RT) 6.50 min (**Figure 5.4A**) which provided the correct m/z (5764.8) for this oligonucleotide by MALDI-TOF-MS. After incubation with **2** at 25 °C for 24 hours in a 4:1 ratio (**2**/DNA), two new minor peaks emerged at RT 12.50 min and RT 14.38 min with the parent AP-1 1, 17I DNA peak still present at RT 6.48 min by 24 hours (**Figure 5B**). The adduct stoichiometry was confirmed by MALDI-TOF-MS as 1:1 **2**/AP-1 1, 17I on the basis of an observed mass of 6320.6 m/z (theoretical mass: 6320.4 m/z) (**Figure 5C**). This result suggested that **2** may form two mono-alkylated adducts at G7 and G12 as an interstrand cross-link between G7 and G12 would be too short to form.

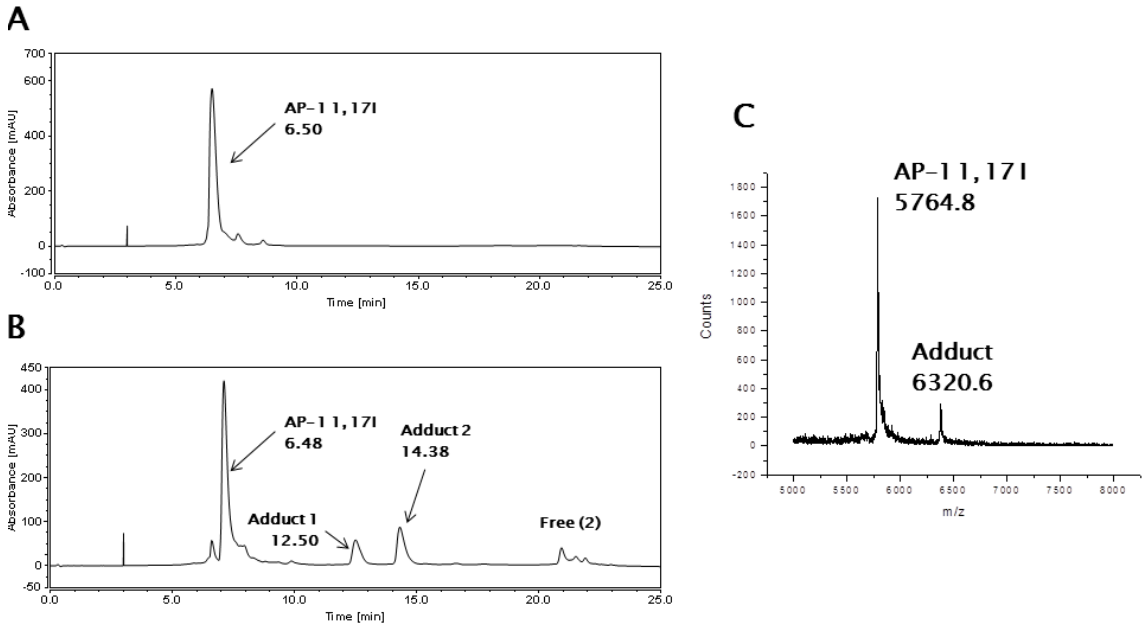


Figure 5.4: **A**, HPLC chromatogram showing the annealed AP-1 1, 17I (**Figure 5.3**) alone at RT 6.50 min; **B**, HPLC chromatogram after incubation of annealed AP-1 1, 17I with **2** for 24 hours, showing approximately 20% conversion to the adduct peaks at RT 12.50 min and RT 14.38 min; **C**, MALDI-TOF spectrum of the adducts at RT 12.50 min and RT 14.38 min from Chromatogram B above. Observed mass of 1:1 **2**/AP-1 1, 17I: 6320.6 m/z (theoretical mass: 6320.4 m/z), observed mass AP-1 1, 17I alone from Chromatogram A: 5764.8 m/z (theoretical mass: 5763.8 m/z).

Next, the interaction of **2** with the AP-1 7, 17I (**Figure 5.3**), with guanines G1 and G12 available for covalent attachment of **2** was studied. Annealed AP-1 7, 17I alone gave a single peak at RT 7.43 min (**Figure 5.5A**) and a m/z of 5764.2 m/z (theoretical mass: 5763.8 m/z). After 24 hours incubation with **2** at 25 °C in a 4:1 ratio (**2**/DNA), a new major peak appeared in the HPLC chromatogram at RT 8.08 min, with ~87% adduct formed (**Figure 5.5B**). The 1:1 stoichiometry of the new peak was confirmed by MALDI-TOF-MS (observed mass: 6321.0 m/z; theoretical mass: 6320.4 m/z) as the **2**/ AP-1 7, 17I adduct (**Figure 5.5C**). This was a surprising result as a PBD had not previously been observed to bond to a terminal guanine (*i.e.*, G1), and G12 was considered to be too close to the TTT-loop to covalently bond to a PBD (**Figure 5.3**). Furthermore, an interstrand cross-link between G1 and G12 would be too long to form. Therefore, our conclusion was that, for AP-1 7, 17I, **2** was bonding to the terminal G1, pointing in the A-ring-3' (*i.e.*, forward) direction (**Figure 5.3**).

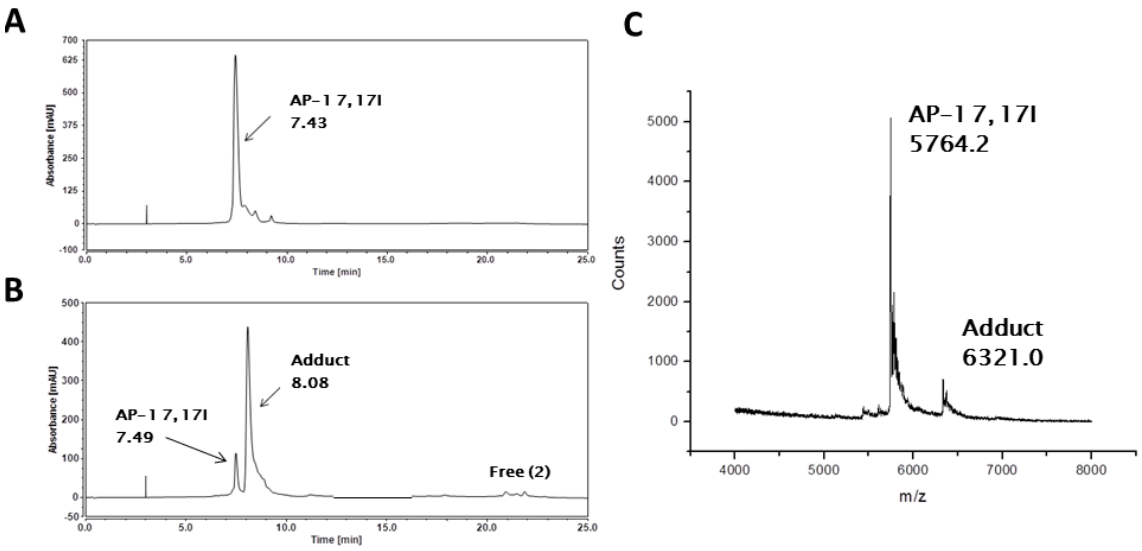


Figure 5.5: **A**, HPLC chromatogram showing the annealed AP-1 7, 17I (**Figure 5.3**) alone at RT 7.43 min; **B**, HPLC chromatogram after incubation of annealed AP-1 7, 17I with **2** for 24 hours, showing approximately 87% conversion to the adduct peak at RT 8.08 min; **C**, MALDI-TOF spectrum of the adduct at RT 8.08 min from Chromatogram B above. Observed mass of 1:1 **2**/AP-1 7, 17I: 6321.0 m/z (theoretical mass: 6320.4 m/z), observed mass AP-1 7, 17I alone from Chromatogram A: 5764.2 m/z (theoretical mass: 5763.8 m/z).

In a similar experiment AP-1 7, 12I (**Figure 5.3**), with guanines G1 and G17 available, was incubated with **2** for 24 hours at 25 °C in a 4:1 ratio (**2**/DNA). Initially, AP-1 7, 12I alone gave a single peak at RT 7.44min (**Figure 5.6A**) with an observed mass of m/z 5764.0 (theoretical mass: m/z 5763.8) (**Figure 5.6C**). After 24 hours incubation, a new major peak at RT 8.08 min was observed (**Figure 5.6B**) with ~89% adduct formation. The stoichiometry of the adduct formed was found to be 1:1 **2**/AP-1 7, 12I by MALDI-TOF-MS analysis (observed mass: 6321.1 m/z, theoretical mass, 6320.4 m/z) (**Figure 5.6C**). In this case, the inosine substitution pattern could have allowed mono-adducts to form at G1 and/or G17, or an interstrand cross-link between the two.

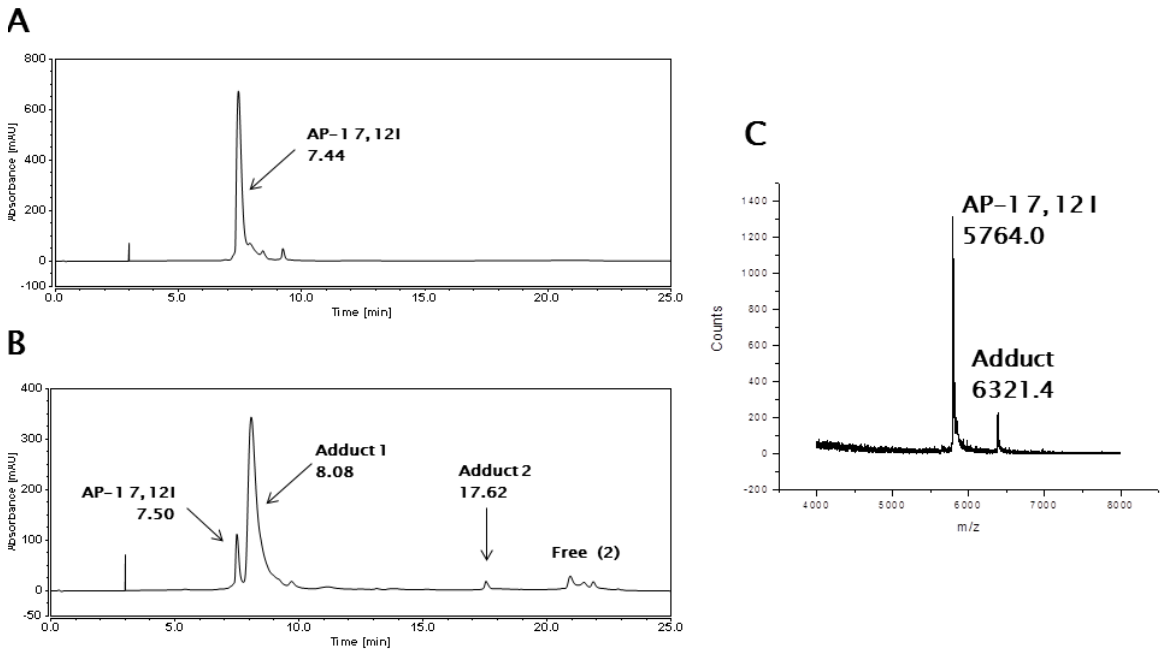


Figure 5.6: **A**, HPLC chromatogram showing the annealed AP-1 7, 12I (**Figure 5.3**) alone at RT 7.44 min; **B**, HPLC chromatogram after incubation of annealed AP-1 7, 12I with **2** for 24 hours, showing approximately 89% conversion to the adduct peak at RT 8.08 min; **C**, MALDI-TOF spectrum of the adduct at RT 8.08 min from Chromatogram B above. Observed mass of 1:1 **2**/AP-1 7, 12I: 6321.4 m/z (theoretical mass: 6320.4 m/z), observed mass AP-1 7, 12I alone from Chromatogram A: 5764.0 m/z (theoretical mass: 5763.8 m/z).

HPLC analysis of annealed AP-1 1, 12I (**Figure 5.3**), with guanines G7 and G17 available showed a single peak at RT 6.88 min (**Figure 5.7A**) which provided the correct m/z (5764.5) for this oligonucleotide by MALDI-TOF-MS. After incubation with **2** at 25 °C for 24 hours in a 4:1 ratio (**2**/DNA), a new major peak emerged at RT 8.41 min with a complete disappearance of the parent AP-1 1, 12I peak by 24 hours (**Figure 5.7B**). The adduct stoichiometry was confirmed by MALDI-TOF-MS as 1:1 **2**/AP-1 1, 12I on the basis of an observed mass of 6321.7 m/z (theoretical mass: 6320.4 m/z) (**Figure 5.7C**). This result suggested that **2** may form a single interstrand cross-link across the G7 and G17 guanines. Furthermore, the extent of reaction after 24 hours was similar to that for the same interstrand cross-link formed in the parent AP-1 sequence.

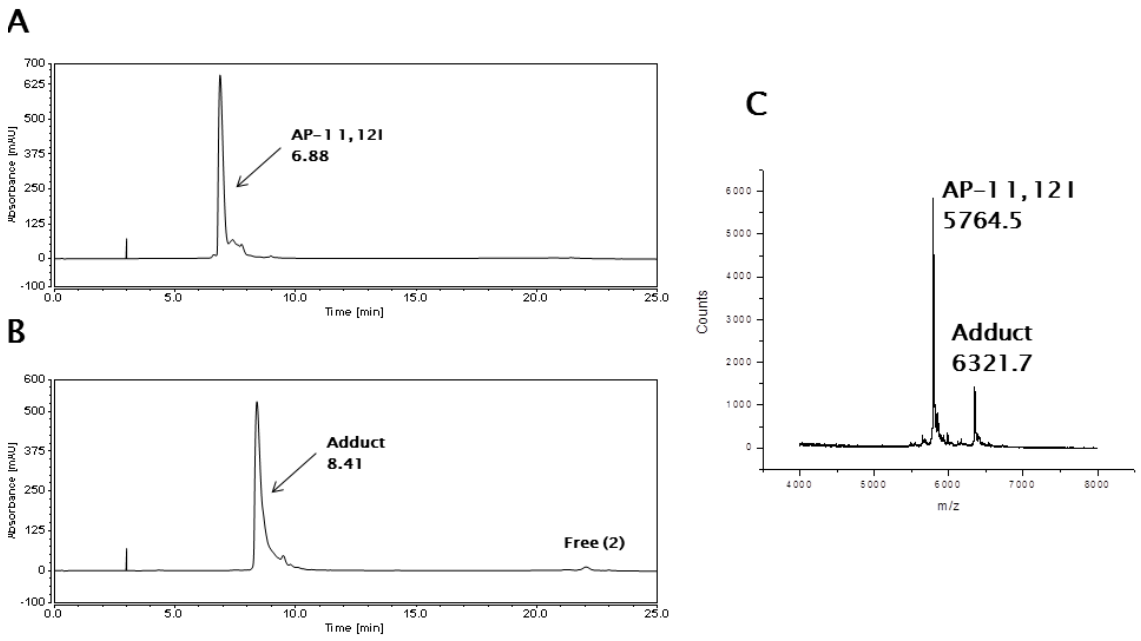


Figure 5.7: **A**, HPLC chromatogram showing the annealed AP-1 1, 12I sequence alone at RT 6.88 min; **B**, HPLC chromatogram after incubation of annealed AP-1 1, 12I with **2** for 24 hours, showing 100% conversion to an adduct at RT 8.41 min with complete loss of the original hairpin peak at RT 6.88 min; **C**, MALDI-TOF spectrum of the adduct (from peak at RT 8.41 min in Chromatogram B above). Observed mass of 1:1 **2**/AP-1 1, 12I adduct: 6321.7 m/z (theoretical mass: 6320.4 m/z), observed mass of AP-1 1, 12I alone from Chromatogram A: 5764.5 m/z (theoretical mass: 5763.8 m/z).

5.3.2 Summary of cross-link formation study of SJG-136

Based on the results obtained from the study on the hairpins shown in **Figure 5.3**, considering that only the extent of reaction of **2** with hairpin AP-1 1, 12I matched its extent of reaction with the parent AP-1 sequence (*i.e.*, 100% adduct formed in 24 hours), this suggested that the G7-G17 interstrand cross-link may be the preferred adduct.

5.4 Interaction of SJG-136 with inosine modified AP-1 hairpin sequences to study mono-alkylation

Based on the observations made for the AP-1 7, 17I sequence (see also Section 5.3.1), the study was continued to investigate the mono-alkylation pattern within the AP-1 sequence by SJG-136. To study the mono-adduct distribution using the developed HPLC/MS methodology, a further four AP-1 hairpins (Hairpin 1-4, **Figure 5.8**) were designed, based on the same AP-1 sequence but which have three of their four guanines replaced with inosines, thus leaving only one reactive guanine (*i.e.*, G1, G7, G12 or G17) in each case, to study mono-alkylation.

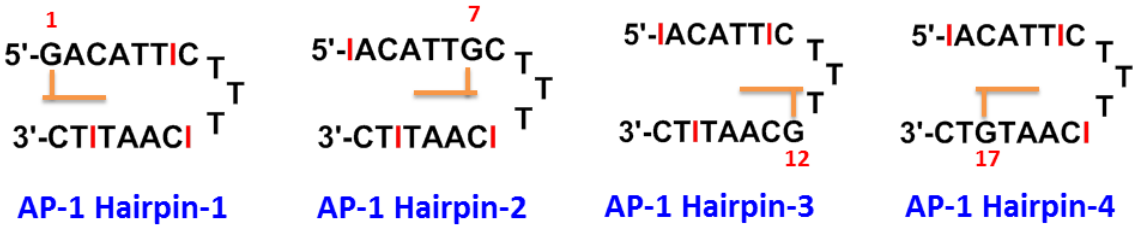


Figure 5.8: Structures of the AP-1 Hairpins 1-4 which are based on the same AP-1 sequence but with three of the four guanines replaced with inosines in each case to study mono-alkylation.

5.4.1 Results and discussion

Working solutions of inosine modified hairpins were prepared according to the procedure described in Section 7.2.1.4. SJG-136 working solutions were prepared as described in Section 7.2.1.9. SJG-136/DNA complexes were prepared according to the protocol described in Section 7.2.1.10. In brief, a SJG-136 working solution of 100 µM was added to a DNA working solution of 25 µM in a 4:1 ratio (SJG-136/DNA). The mixture was mixed and incubated for 24 hours at 25 °C. Following the incubation, the samples were subjected to ion pair RP-HPLC/MALDI-TOF-MS analysis as described in Sections 7.2.2 and 7.2.4 in detail.

HPLC analysis of annealed AP-1 Hairpin-1 (**Figure 5.8**), with only the terminal guanine available gave a single peak at RT 7.64 min (**Figure 5.9A**) which provided the correct

m/z (5748.3) for this oligonucleotide by MALDI-TOF-MS. Following a 24 hour incubation with **2** at 25 °C in a 4:1 molar ratio (**2**/DNA), a new peak appeared in the HPLC chromatogram at RT 14.03 min (**Figure 5.9B**). The adduct stoichiometry was confirmed by MALDI-TOF-MS as 1:1 **2**/AP-1 Hairpin-1 based on an observed mass of 6305.1 m/z (theoretical mass: 6305.41 m/z) (**Figure 5.9C**). Although the reaction had not gone to completion by 24 hours (approximately 50%), this result confirmed that a PBD unit could bond to a terminal G1 base, with the body of the PBD spanning two rather than three base pairs. The most likely explanation for this is that one PBD unit of **2** bonds to the terminal 5'-guanine with its C-ring protruding just outside of the minor groove, and with the second PBD unit laying in the minor groove (pointing toward the loop) and making non-covalent interactions without forming a second cross-linking covalent bond. This structure was supported by a molecular dynamics simulation (see also Section 5.9) which suggested that a stable adduct should form with **2** arranged in this way. The fact that adduct formation did not go to completion after 24 hours is also consistent with a mono-alkylation event in which the covalently bonded PBD is interacting with two rather than the usual three base pairs at the alkylation site.

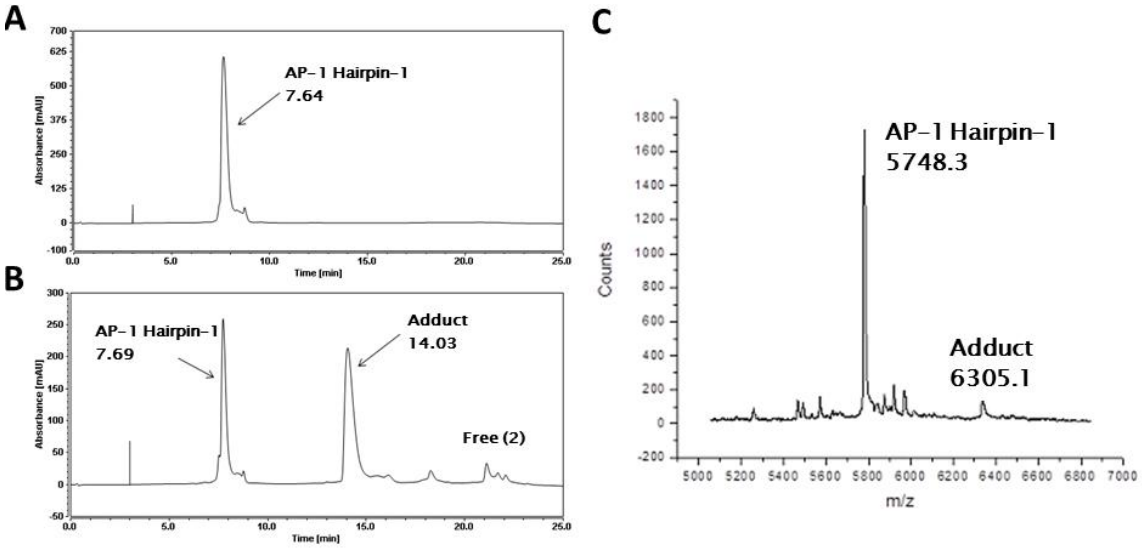


Figure 5.9: **A**, HPLC chromatogram showing the annealed AP-1 Hairpin-1 (**Figure 5.8**) alone at RT 7.64 min; **B**, HPLC chromatogram after incubation of annealed AP-1 Hairpin-1 with **2** for 24 hours, showing approximately 50% conversion to the adduct peak at RT 14.03 min; **C**, MALDI-TOF spectrum of the adduct at RT 14.03 min from Chromatogram B above. Observed mass of 1:1 **2**/AP-1 Haipin-1 adduct: 6305.1 m/z (theoretical mass: 6305.41 m/z), observed mass of AP-1 Hairpin-1 alone from Chromatogram A: 5748.3 m/z (theoretical mass: 5748.8 m/z).

Next, the interaction of **2** with the AP-1 Hairpin-2 (**Figure 5.8**) was studied. Annealed AP-1 Hairpin-2 alone resulted in the appearance of a single peak at RT 8.01 min (**Figure 5.10A**), which was identified by MALDI-TOF-MS (m/z 5749.2). After 24 hours incubation with **2** at 25 °C in a 4:1 molar ratio (**2**/DNA), a new peak emerged in the HPLC chromatogram at RT 9.24 min, although reaction had not gone to completion (**Figure 5.10B**). The 1:1 stoichiometry of the adduct (**2**/DNA) was confirmed by MALDI-TOF-MS (observed mass: 6306.1 m/z, theoretical mass: 6305.41 m/z) (**Figure 5.10C**), and the structure was concluded to be the mono-alkylated adduct with **2** bonded to G7 and with the bulk of the PBD Dimer lying in the minor groove pointing away from the loop. The conversion within 24 hours (approximately 50%) was assumed to reflect the lower stability of a mono- rather than a cross-linked adduct.

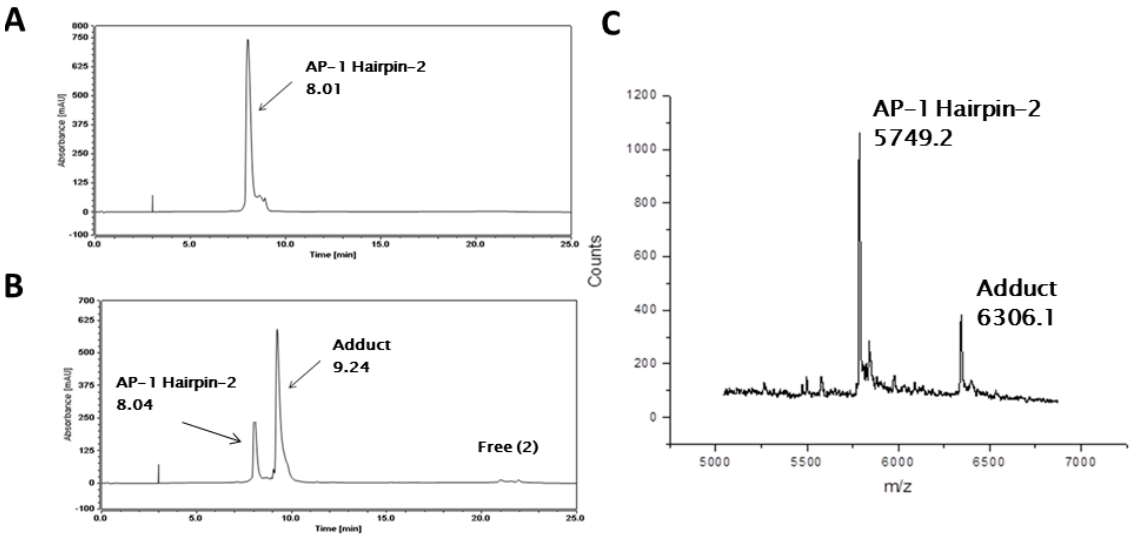


Figure 5.10: HPLC chromatograms: **A**, Annealed AP-1 Hairpin-2 at RT 8.01 min; **B**, Annealed AP-1 Hairpin-2 after incubating with **2** for 24 hours, showing one new major adduct peak at RT 9.24 min with reaction not complete after 24 hours; **C**, MALDI-TOF spectrum of the adduct (from peak at RT 9.24 min in Chromatogram B above). Observed mass of AP-1 Hairpin-2: 5749.2 m/z (theoretical mass: 5748.8 m/z); Observed mass of **2**/AP-1 Hairpin-2 adduct: 6306.1 m/z (theoretical mass: 6305.41 m/z).

Similar results were obtained for AP-1 Hairpin-3 (**Figure 5.8**), with only G12 available for covalent interaction with **2**. HPLC analysis of annealed AP-1 Hairpin-3 alone provided a single peak at RT 8.14 min (**Figure 5.11A**) identified by MALDI-TOF-MS (m/z 5748.60). After incubation with **2** for 24 hours at 25 °C in a 4:1 molar ratio (**2**/DNA), a new peak was observed at RT 9.07 min (**Figure 5.11B**), identified as the 1:1 **2**/AP-1 Hairpin-3 adduct by MALDI-TOF-MS (observed mass: 6305.0 m/z, theoretical mass: 6305.41 m/z) (**Figure 5.11C**). Again, the reaction was approximately 80% complete, consistent with mono-adduct formation. This result was also surprising due to the proximity of G12 to the loop structure. Molecular Dynamics Simulations (see also Section 5.9) suggested that, although located close to the TTT-loop, a **2**:G12 mono-alkylated adduct may be feasible**,** with the second PBD of the dimer orientated away from the loop and forming non-covalent interactions in the minor groove.

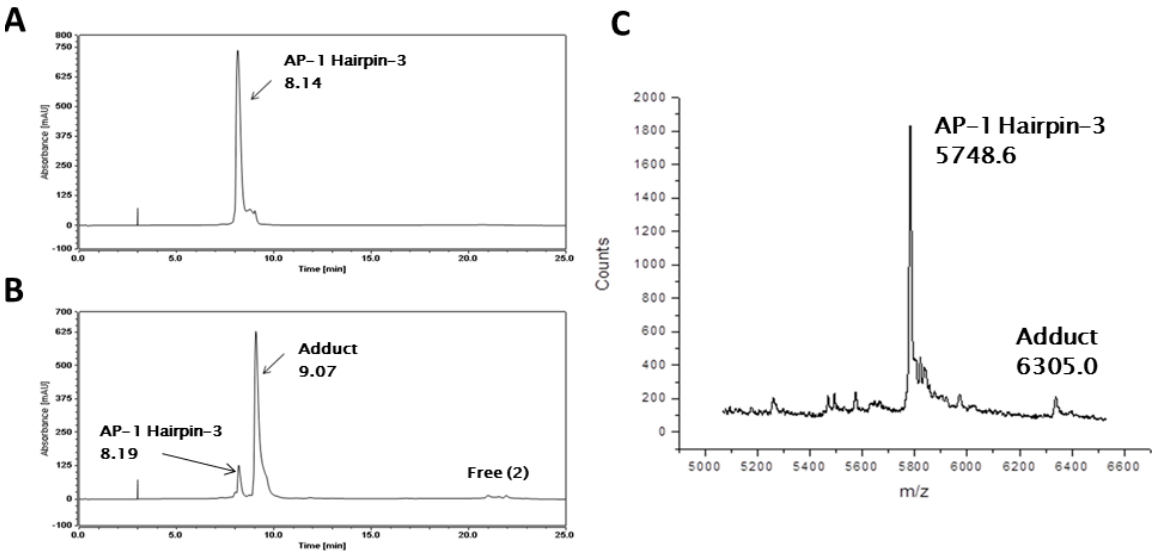


Figure 5.11: HPLC chromatograms: **A**, Annealed AP-1 Hairpin-3 at RT 8.14 min; **B**, Annealed AP-1 Hairpin-3 after incubating with **2** for 24 hours, showing one new major adduct peak at RT 9.07 min with reaction not complete after 24 hours; **C**, MALDI-TOF spectrum of the adduct (from peak at RT 9.07 min in Chromatogram B above). Observed mass of AP-1 Hairpin-3: 5748.6 m/z (theoretical mass: 5748.8 m/z); Observed mass of **2**/AP-1 Hairpin-3 adduct: 6305.0 m/z (theoretical mass: 6305.41 m/z).

Lastly, **2** was incubated with the AP-1 Hairpin-4 (**Figure 5.8**) for 24 hours at 25 °C in a 4:1 molar ratio (**2**/DNA). The HPLC chromatogram of annealed AP-1 Hairpin-4 alone provided a distinct peak at RT 8.03 min (**Figure 5.12A**) identified by MALDI-TOF-MS. After 24 hours incubation with **2**, a new peak was observed at 8.86 min (**Figure 5.12B**) with reaction proceeding to approximately 70% completion. MALDI-TOF-MS analysis confirmed the new peak at RT 8.86 min to be the 1:1 **2**/AP-1 Hairpin-4 adduct, with an observed mass of 6305.6 m/z (theoretical mass: 6305.41 m/z) (**Figure 5.12C**). For this adduct there should be no steric issues, with one of the PBD units bonding to G17 and the bulk of the molecule lying in the minor groove pointing toward the loop.

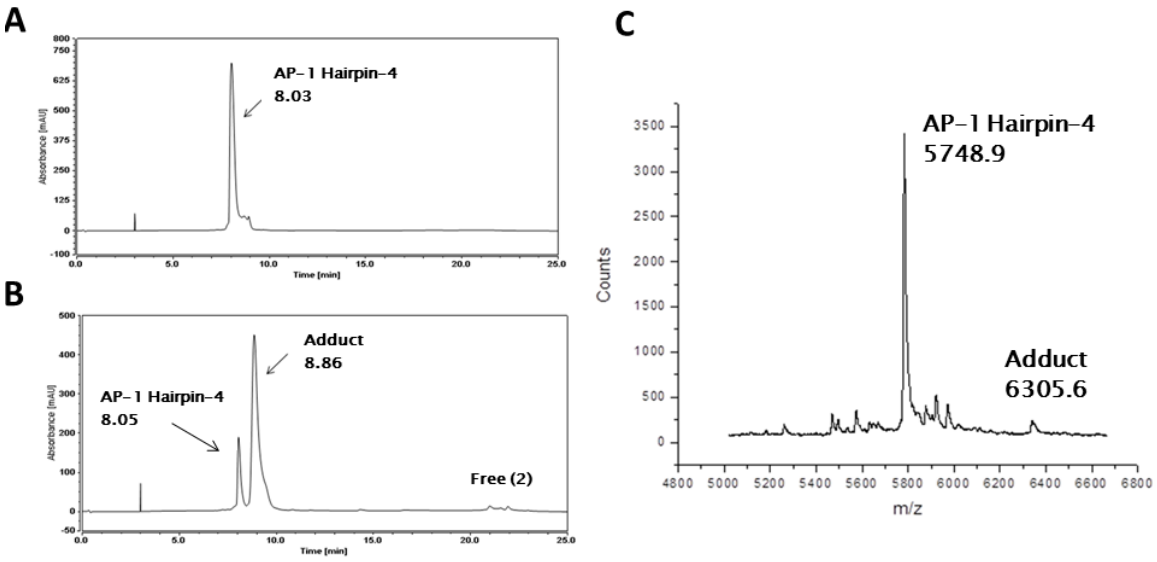


Figure 5.12: HPLC chromatograms: **A**, Annealed AP-1 Hairpin-4 at RT 8.03 min; **B**, Annealed AP-1 Hairpin-4 after incubating with **2** for 24 hours, showing one new major adduct peak at RT 8.86 min with reaction not complete after 24 hours; **C**, MALDI-TOF spectrum of the adduct (from peak at RT 8.86 min in Chromatogram B above). Observed mass of AP-1 Hairpin-4: 5748.9 m/z (theoretical mass: 5748.8 m/z); Observed mass of **2**/AP-1 Hairpin-4 adduct: 6305.6 m/z (theoretical mass: 6305.41 m/z).

5.4.2 Summary of mono-alkylation study of SJG-136

In summary, the results from the experiments with the AP-1 Hairpins 1-4 demonstrated that all four guanines (*i.e.*, G1, G7, G12 or G17) are reactive toward **2**, and are capable of forming mono-alkylated adducts. The reactivity of G1 was particularly surprising, as it means that the covalently bonded PBD must span only two base pairs, a phenomenon not previously observed and not previously thought to be possible. Previous studies on the DNA binding characteristics of PBDs have suggested that a minimum of three consecutive DNA bases with a central guanine (*i.e.*, Pu-G-Pu) are required for covalent attachment[297].

5.5 Interaction of anthramycin (1) and GWL-78 (3) with the inosine-modified AP-1 Hairpin-1 sequence

To determine whether the size of the PBD molecule might affect the mode of binding to the AP-1 hairpins, similar experiments were carried out with the PBD monomers anthramycin (**1**) and the PBD C8-conjugate GWL-78 (**3**) (**Figure 5.1**).

5.5.1 Results and discussion

Working solutions of inosine modified AP-1 Hairpin-1 sequence were prepared according to the protocol described in Section 7.2.1.4. Anthramycin and GWL-78 working solutions were prepared as described in Sections 7.2.1.6 and 7.2.1.7. PBD/DNA complexes were prepared according to the procedure described in Section 7.2.1.10. In brief, a PBD working solution of 100 µM was added to a DNA working solution of 25 µM in a 4:1 ratio (PBD/DNA). The mixture was agitated and incubated for 24 hours at 25 °C. Following the incubation, the samples were subjected to ion pair RP-HPLC/MALDI-TOF-MS analysis as described in Sections 7.2.2 and 7.2.4 in detail.

First, **1** was incubated with the AP-1 Hairpin-1 (with only the 5'-terminal guanine available for covalent binding) (**Figure 5.8**) in a 4:1 ratio (**1**/DNA) for 24 hours at 25 °C. However, no changes were observed in the HPLC chromatogram, with the unreacted AP-1 Hairpin-1 peak at RT 7.90 min remaining the major peak (**Figures 5.13A and 5.13B**) as confirmed by MALDI-TOF-MS (observed mass: 5748.0 m/z, theoretical mass: 5748.8 m/z) (**Figure 5.13C**). This suggested that substitution at the C8-position, as in SJG-136 (**2**) may have a significant influence on the binding properties of a PBD, enhancing interaction of the molecule in the DNA minor groove through non-covalent interactions (*e.g.*, van der Waals, hydrogen bonds and/or electrostatic interactions). The next step was to study GWL-78 (**3**), a synthetic PBD monomeric conjugate substituted at the C8 position with a *bis*-pyrrole side-chain that allows the molecule to span a total of five to six DNA base pairs. It was anticipated that this molecule may behave more like SJG-136 (**2**) than anthramycin (**1**) due to the large non-covalent DNA-binding moiety at the C8-position. As predicted, when **3** was incubated with the AP-1 Hairpin-1 (RT 8.69 min, **Figure 5.13D**) in a 4:1 ratio (**3**/DNA) for 24 hours at 25 °C, a new peak was observed at

11.13 min (**Figure 5.13E**), although the extent of reaction was much lower compared to **2**. The new peak was identified as the 1:1 **3**/AP-1 Hairpin-1 adduct by MALDI-TOF-MS (observed mass: 6338.6 m/z, theoretical mass: 6339.4 m/z) (**Figure 5.13F**).

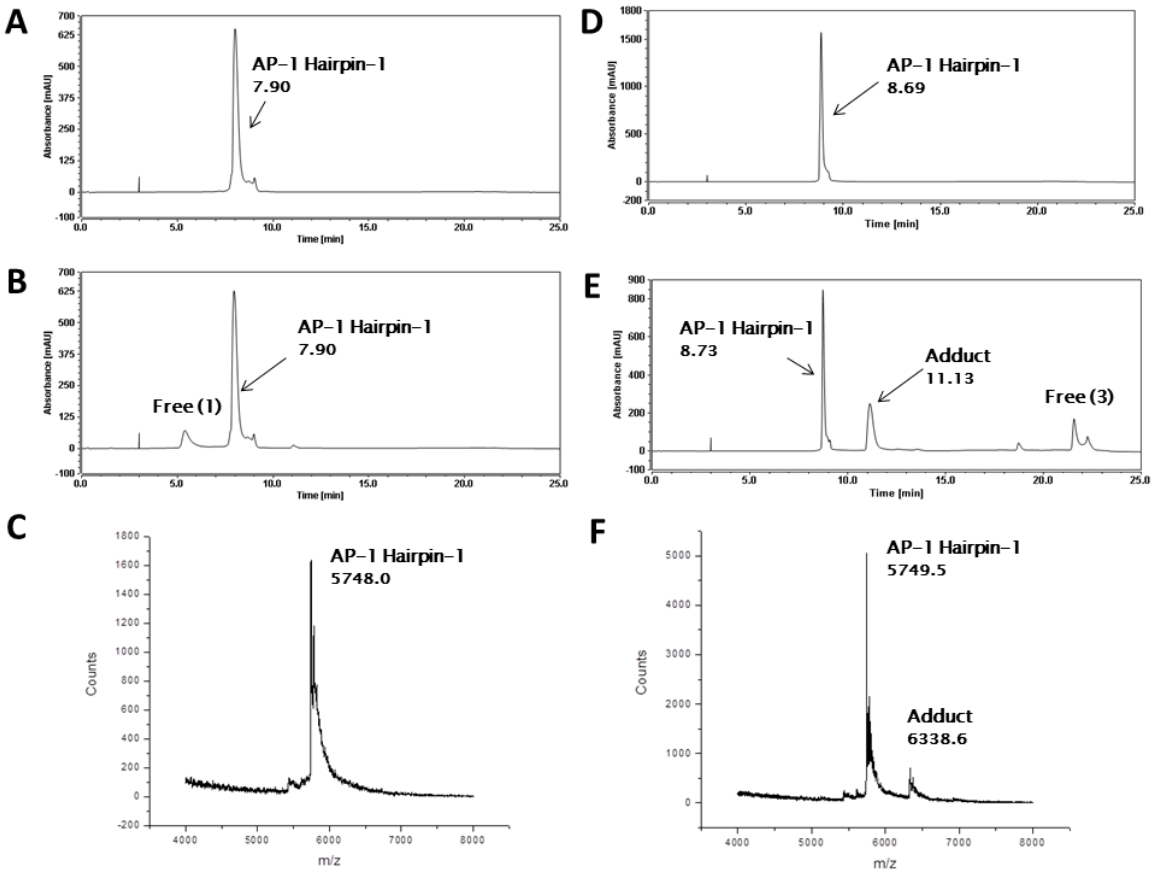


Figure 5.13: **A**, HPLC chromatogram showing annealed AP-1 Hairpin-1 at RT 7.90 min; **B**, Annealed AP-1 Hairpin-1 after incubating with **1** for 24 hours showing no adduct formation after 24 hours; **C**, MALDI-TOF spectrum of **1** with the AP-1 Hairpin-1 confirming that no adduct had formed (AP-1 Hairpin-1 observed mass: 5748.0 m/z, theoretical mass: 5748.8 m/z); **D**, Annealed AP-1 Hairpin-1 at RT 8.69 min; **E**, Annealed AP-1 Hairpin-1 after incubating with **3** for 24 hours showing the appearance of an adduct peak at RT 11.13 min with approximately 23% complete reaction after 24 hours; **F**, MALDI-TOF spectrum of **3**/AP-1 Hairpin-1 confirming the identity of adduct formation (AP-1 Hairpin-1 observed mass: 5749.5 m/z, theoretical mass: 5748.8 m/z, AP-1 Hairpin-1 adduct observed mass: 6338.6 m/z, theoretical mass: 6339.4 m/z). The observed differences in the RT of AP-1 Hairpin-1 alone (*i.e.*, **Figure 5.13A** and **Figure 5.13D**) occurred due to mobile phase changes.

5.5.2 Summary of the interaction of anthramycin and GWL-78 with the AP-1 Hairpin-1 sequence

From these results it was concluded that covalent bonding of a PBD to a terminal guanine may only occur when there is sufficient molecular bulk at the C8-position (*e.g.*, another PBD unit in the case of SJG-136, or a *bis*-pyrrole unit in the case of GWL-78) to provide stabilisation to the adduct by interacting non-covalently in the adjacent minor groove.

5.6 Interaction of anthramycin (1), SJG-136 (2) and GWL-78 (3) with the AP-1 1 Guanine sequence

In order to rule out whether these new observations of a PBD monomer (**3**) and dimer (**2**) bonding to a terminal guanine might be due to a change in DNA conformation caused by the insertion of inosine bases, the interaction of **1**, **2**, and **3** with AP-1 1 Guanine (1G) (**Figure 5.14**) was studied. This hairpin contains a 5'-terminal guanine without any inosine modifications. Instead, G7, G12 and G17 were replaced with adenine bases, and their corresponding cytosines with thymines.

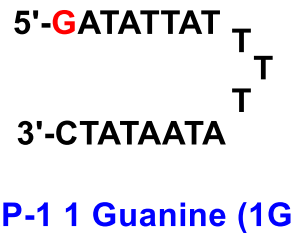


Figure 5.14: AP-1 1 Guanine hairpin sequence which is based on the same AP-1 sequence but with three of the four guanine bases (except the 5'-terminal-guanine) mutated to an A.

5.6.1 Results and discussion

Working solutions of inosine modified AP-1 1G 1 sequence were prepared according to the protocol described in Section 7.2.1.4. Anthramycin, GWL-78 and SJG-136 working solutions were prepared as described in Sections 7.2.1.6, 7.2.1.7 and 7.2.1.9. PBD/DNA complexes were prepared according to the procedure described in Section 7.2.1.10. Briefly, a PBD working solution of 100 µM was added to a DNA working solution of 25

μM in a 4:1 ratio (PBD/DNA). The mixture was agitated and incubated for 24 hours at 25 °C. Following the incubation, the samples were subjected to ion pair RP-HPLC/MALDI-TOF-MS analysis as described in Sections 7.2.2 and 7.2.4 in detail.

The AP-1 1G hairpin sequence (**Figure 5.14**) was incubated with **1**, **2**, and **3** in a 4:1 ratio (PBD/DNA) for 24 hours at 25 °C, and immediately subjected to HPLC and MALDI-TOF-MS analysis. AP-1 1G hairpin alone gave a single peak at RT 9.56 min (**Figure 5.15A**) identified by MALDI-TOF-MS. After reaction with **2**, one new major peak at RT 14.19 min appeared along with a minor peak at RT 18.40 min (**Figure 5.15B**). The stoichiometry of the main new peak as the 1:1 **2**/AP-1 1G hairpin was confirmed by MALDI-TOF-MS analysis with an observed mass of 6347.9 m/z (theoretical mass: 6347.5 m/z) (**Figure 5.15E**). Incubation of **3** with the AP-1 1G hairpin gave one new major peak at RT 12.07 min and a minor peak at RT 18.81 min (**Figure 5.15C**). Reaction was not complete after 24 hours with some AP-1 1G hairpin remaining (**Figure 5.15C**). The stoichiometry of the main adduct formed was confirmed by MALDI-TOF-MS analysis as the 1:1 **3**/AP-1 1G hairpin with a mass of 6381.2 m/z (theoretical mass: 6381.5 m/z) (**Figure 5.15F**). The identity of the minor peaks at RT 18.40 min (for **2**) and RT 18.81 min (for **3**) could not be confirmed. Lastly, HPLC analysis of the interaction of **1** with the AP-1 1G hairpin provided no changes in the chromatogram (**Figure 8D**), with MALDI-TOF-MS (**Figure 5.15G**) confirming that no adduct had formed.

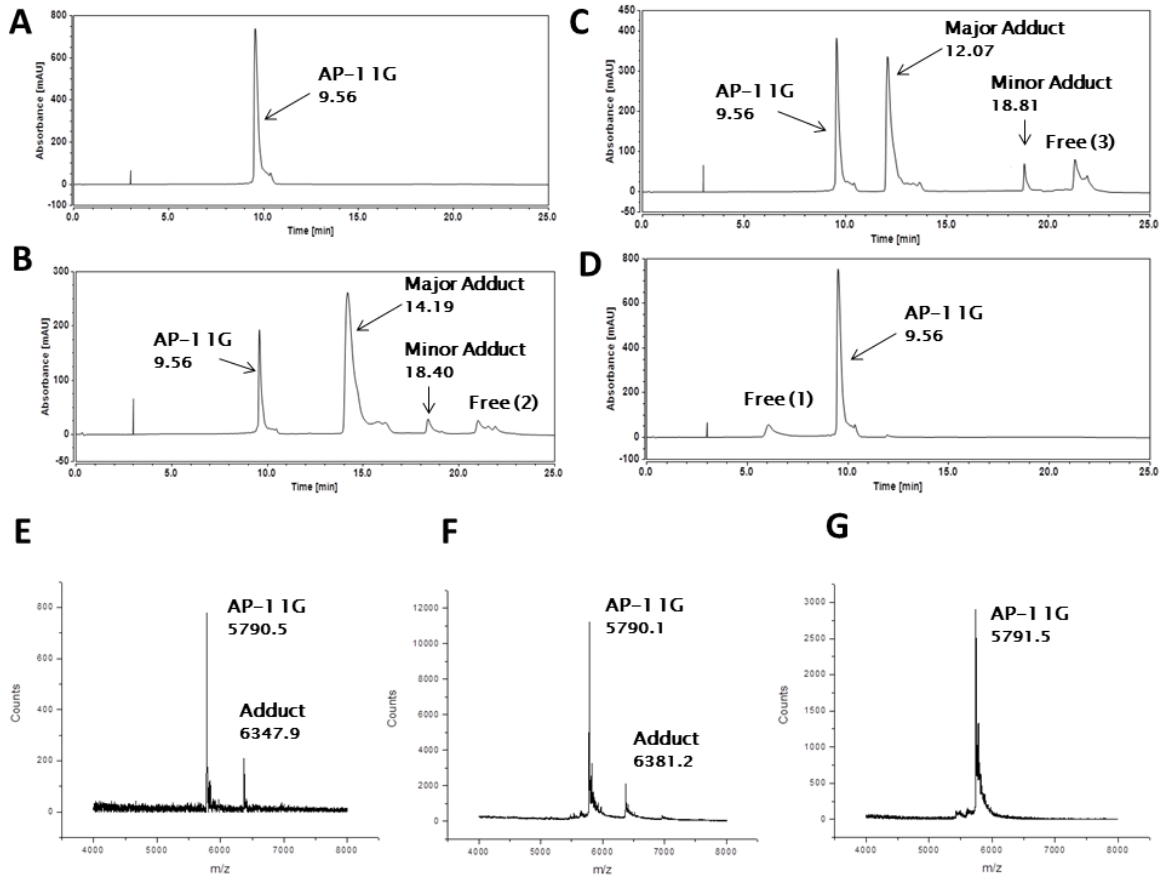


Figure 5.15: HPLC chromatograms: **A**, Annealed AP-1 1G hairpin at RT 9.56 min; **B**, Annealed AP-1 1G hairpin after incubating with **2** for 24 hours with the appearance of one new major adduct peak at RT 14.19 and a minor peak at RT 18.40 min; **C**, Annealed AP-1 1G hairpin after incubating with **3** for 24 hours with the appearance of one new major adduct peak at RT 12.07 and a minor peak at RT 18.81 min; **D**, Annealed AP-1 1G hairpin after incubating with **1** for 24 hours showing no adduct formation after 24 hours; **E**, MALDI-TOF spectrum of **2** with the AP-1 1G hairpin confirming the identity of adduct formation, AP-1 1G hairpin observed mass: 5790.5 m/z (theoretical mass: 5790.9 m/z), **2**/ AP-1 1G hairpin adduct observed mass: 6347.9 m/z (theoretical mass: 6347.5 m/z); **F**, **3** with the AP-1 1G hairpin confirming the stoichiometry of adduct formation, AP-1 1G hairpin observed mass: 5790.1 m/z (theoretical mass: 5790.9 m/z), **3**/AP-1 1G hairpin adduct observed mass: 6381.2 m/z (theoretical mass: 6381.5 m/z); **G**, **1** with the AP-1 1G hairpin confirming no adduct formation, AP-1 1G hairpin observed mass: 5791.5 m/z (theoretical mass: 5790.9 m/z).

5.6.2 Summary of the interaction of anthramycin, GWL-78 and SJG-136 with AP-1 1G hairpin sequence

In summary, these results supported the observations made with the AP-1 Hairpin-1, and confirmed that the inosine mutations had not altered the reactivity of the DNA toward the PBDs in any way.

5.7 Interaction of SJG-136 (2) with double-stranded inosine-modified AP-1 sequences

In order to explore whether the observations made within the inosine-modified AP-1 hairpins were a consequence of their hairpin structure, a similar series of experiments was carried out on the sequence-related DNA duplexes shown in **Figure 5.16**. As with the AP-1 hairpin oligonucleotides 1-4 (**Figure 5.8**), three of the four guanines available for covalent attachment of a PBD were replaced with non-nucleophilic inosines resulting in only one guanine available for covalent bonding in each case. Each duplex oligonucleotide had a length of 8 base pairs, as it has been previously demonstrated by this laboratory (*unpublished data*) that a minimum of 7 base pairs is required to ensure a DNA minor-groove environment.

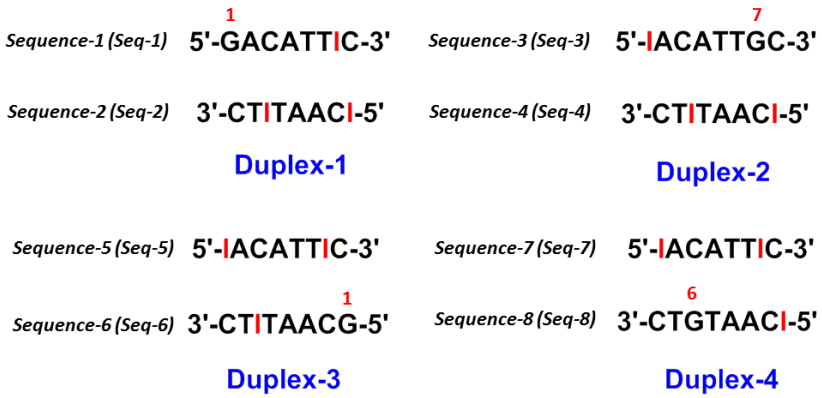


Figure 5.16: Structures of the inosine-modified DNA duplexes used in the study. Duplexes 1-4 are analogous to AP-1 Hairpins 1-4 (**Figure 5.8**) in that three of the four guanines have been replaced with inosines. The duplexes are 8 base pairs in length.

5.7.1 Results and discussion

Working solutions of inosine modified duplex oligonucleotides were prepared according to the procedure described in Section 7.2.1.5. SJG-136 working solutions were prepared as described in Section 7.2.1.9. SJG-136/DNA complexes were prepared according to the procedure described in Section 9.2.1.10. Briefly, a SJG-136 working solution of 100 µM was added to a DNA working solution of 25 µM in a 4:1 ratio (PBD/DNA). The mixture was mixed and incubated for 24 hours at 25 °C. Following the incubation, the samples were subjected to ion pair RP-HPLC/MALDI-TOF-MS analysis as described in Sections 7.2.2 and 7.2.4 in detail.

Annealed Duplex-1 (**Figure 5.16**) gave two peaks in the HPLC chromatogram at RT 7.16 min (Seq-2) and RT 7.46 min (Seq-1) (**Figure 5.17A**) identified as the single strands of Duplex-1 by MALDI-TOF-MS with detected masses: Seq-2, 2378.8 m/z and Seq-1, 2395.2 m/z (theoretical masses: Seq-2, 2379.6 m/z and Seq-1, 2394.6 m/z) (**Figure 5.17C**). This result was in accord with previous reports that double-stranded oligonucleotides of this length denature under these HPLC conditions[157]. Duplex-1 was then incubated with **2** in a 4:1 ratio (**2**/DNA) for 24 hours at 25 °C, and subjected to HPLC analysis. A new minor peak appeared at RT 19.30 min (**Figure 5.17B**) suggesting that an adduct had formed. Reaction was incomplete after 24 hours, with peaks corresponding to Seq-2 (RT 7.20 min) and Seq-1 (RT 7.52 min) remaining (**Figure 5.17B**). Subsequent MALDI-TOF-MS analysis confirmed the presence of a 1:1 **2**/Seq-1 adduct with an observed mass of 2951.8 m/z (theoretical mass: 2951.21 m/z) (**Figure 5.17C**). It is noteworthy that in a similar manner to the denaturation that occurs during the HPLC process, under MALDI-TOF-MS conditions the duplex denatures into single strands with the PBD still attached to the guanine-containing strand. This observation was consistent with the data obtained for AP-1 Hairpin-1, as Duplex-1 contains only one available guanine at its 5'-terminus. This further confirmed the ability of **2** to interact with a terminal guanine.

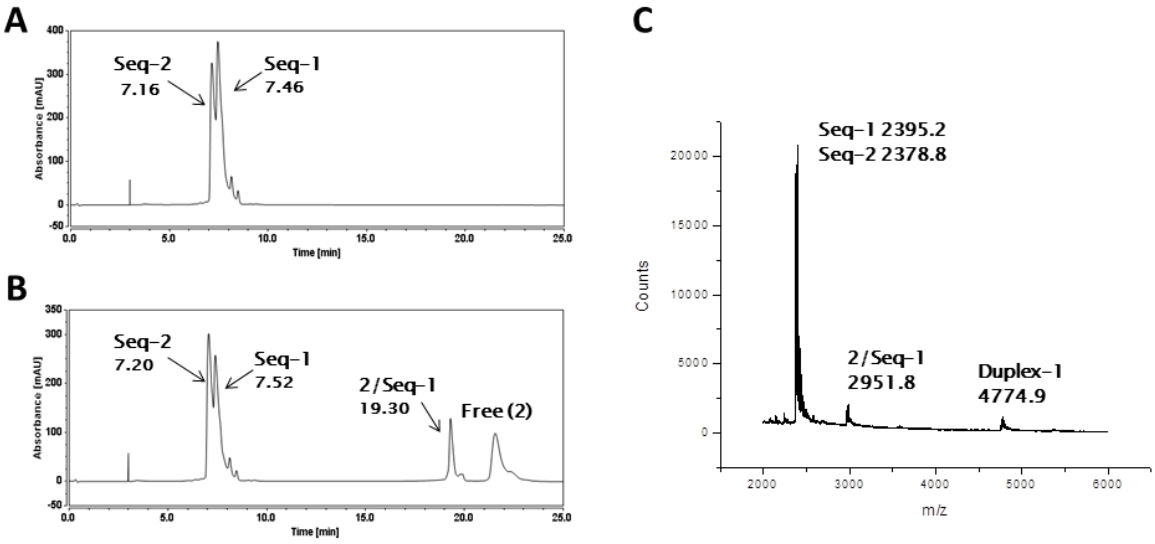


Figure 5.17: HPLC chromatograms; **A**, Annealed Duplex-1 at RT 7.16 min (Seq-2) and RT 7.46 min (Seq-1), showing that the duplex had denatured under the HPLC conditions. **B**, Annealed Duplex-1 after incubating with **2** for 24 hours showing the appearance of one new adduct peak at RT 19.30 min (**2**/Seq-1), with the reaction not complete after 24 hours (~13% adduct conversion); **C**, MALDI-TOF spectrum of **2** with Duplex-1, confirming the stoichiometry of the adduct formed; single strands of Duplex-1, observed masses: Seq-2, 2378.8 m/z and Seq-1, 2395.2 m/z (theoretical mass: Seq-2 2379.6 m/z and Seq-1 2394.6 m/z), and Duplex-1 observed mass: 4774.9 m/z (theoretical mass: 4774.2 m/z), **2**/Seq-1 adduct observed mass: 2951.8 m/z (theoretical mass: 2951.21 m/z). The insignificant differences in the RT for the single-stranded oligonucleotides in chromatograms A and B are due to HPLC run-time variations.

Next, Duplex-2 (**Figure 5.16**) was studied. The resulting chromatogram provided two peaks at RT 7.13 min and RT 7.27 min corresponding to Seq-4 and Seq-3, respectively (**Figure 5.18A**). Their stoichiometry was confirmed by MALDI-TOF-MS with detected masses of m/z 2378.9 and 2394.5 (theoretical masses: Seq-4, 2379.6 m/z and Seq-3, 2394.6 m/z) (**Figure 5.18C**). After 24 hours incubation with **2** at 25 °C in a 4:1 molar ratio (**2**/DNA), a new minor peak appeared at RT 19.30 min (**Figure 5.18B**). Reaction was incomplete after 24 hours (~14% adduct formation) with the two major peaks at RT 7.16 min (Seq-4) and RT 7.32 min (Seq-3) remaining (**Figure 5.18B**). The stoichiometry of the adduct formed was confirmed by MALDI-TOF-MS analysis as the 1:1 **2**/Seq-3 adduct with a detected mass of 2952.0 m/z (theoretical mass: 2951.21 m/z) (**Figure 5.18C**). This result was similar to that obtained for AP-1 Hairpin-2, and the slow extent

of reaction may be explained by the least preferred A-ring-5' oriented adduct that must form, as the guanine is close to the 3'-end of the duplex.

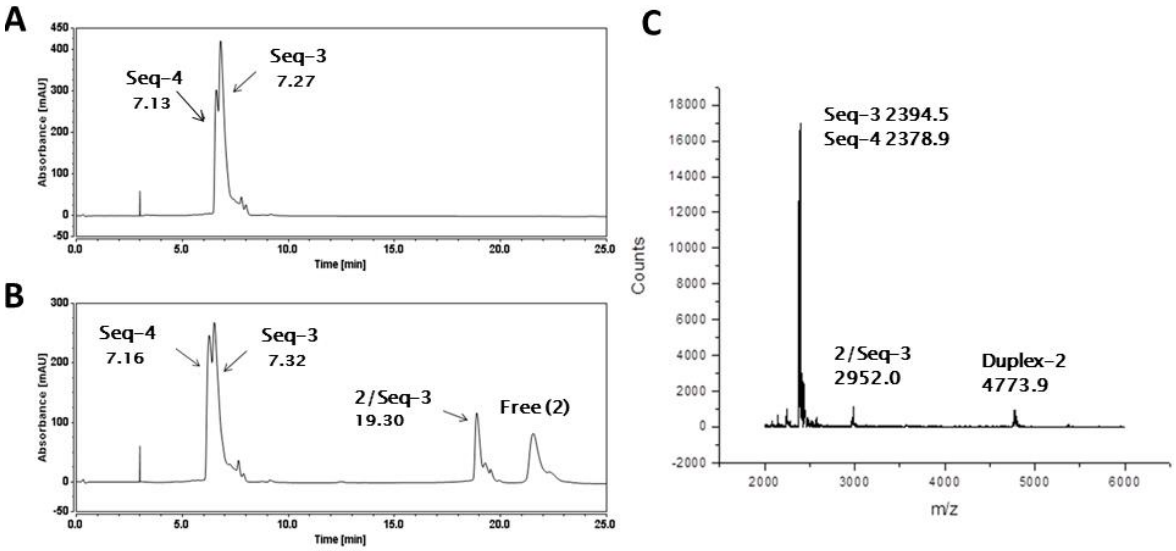


Figure 5.18: HPLC chromatograms; **A**, Annealed Duplex-2 at RT 7.13 min (Seq-4) and RT 7.27 min (Seq-3); **B**, Annealed Duplex-2 after incubating with **2** for 24 hours showing a new adduct peak at RT 19.30 min (**2**/Seq-3), with the reaction not complete after 24 hours; **C**, MALDI-TOF spectrum of **2** with Duplex-2, confirming the stoichiometry of the adduct formed; single strands of Duplex-2, observed masses: Seq-4, 2378.9 m/z and Seq-3, 2394.5 m/z (theoretical mass: Seq-4 2379.6 m/z and Seq-3 2394.6 m/z), and Duplex-2 observed mass: 4773.9 m/z (theoretical mass: 4774.2 m/z), **2**/Seq-3 adduct observed mass: 2952.0 m/z (theoretical mass: 2951.21 m/z).

HPLC analysis of Duplex-3 (**Figure 5.16**) provided two peaks at RT 6.86 min (Seq-6) and RT 7.40 min (Seq-5) (**Figure 5.19A**) identified by MALDI-TOF-MS as the component strands with observed masses of 2378.9 m/z (Seq-5) and 2394.2 m/z (Seq-6) (theoretical masses: Seq-5, 2379.6 m/z and Seq-6, 2394.6 m/z) (**Figure 5.19C**). Incubation for 24 hours at 25 °C with **2** in a 4:1 ratio (**2**/DNA) provided a major new peak at RT 18.30 min, the stoichiometry of which was confirmed by MALDI-TOF-MS as the 1:1 **2**/Seq-6 adduct with an observed mass of 2951.5 m/z (theoretical mass: 2951.21 m/z) (**Figure 5.19C**). However, the reaction was incomplete after 24 hours with ~43% adduct formation (**Figure 5.19B**). Interestingly, the extent of reaction of **2** with Duplex-3 was greater than that for Duplex-1 or Duplex-2, providing further evidence that a PBD can react at a terminal guanine base.

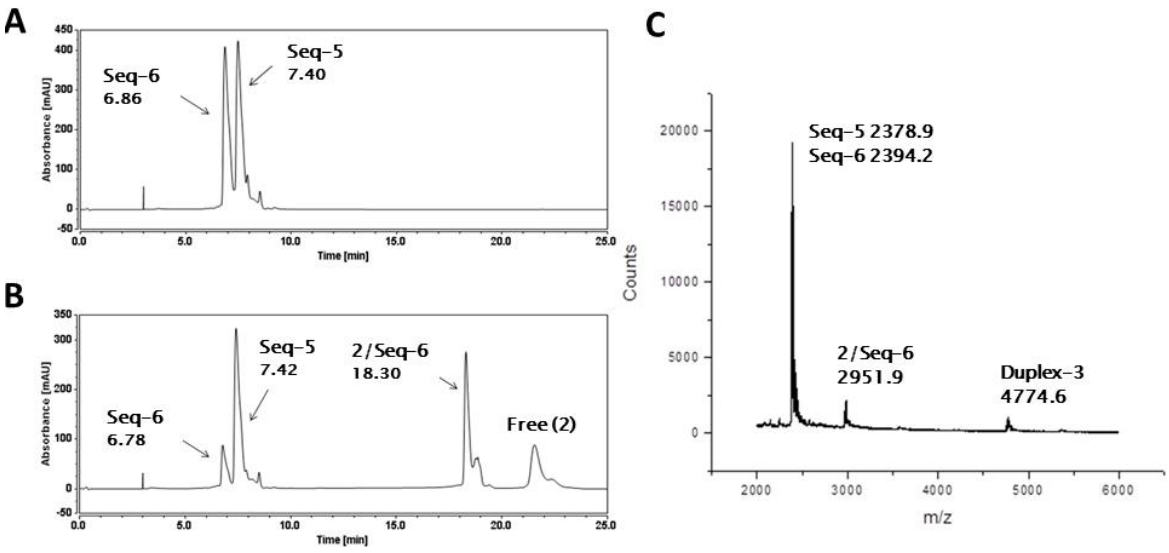


Figure 5.19: HPLC chromatograms; **A**, Annealed Duplex-3 at RT 6.86 min (Seq-6) and RT 7.40 min (Seq-5); **B**, Annealed Duplex-3 after incubating with **2** for 24 hours showing a new adduct peak at RT 18.30 min (**2**/Seq-6), with the reaction not complete after 24 hours; **C**, MALDI-TOF spectrum of **2** with Duplex-3, confirming the stoichiometry of the adduct formed; single strands of Duplex-3, observed masses: Seq-5, 2378.9 m/z and Seq-6, 2394.2 m/z (theoretical mass: Seq-5 2379.6 m/z and Seq-6 2394.6 m/z), and Duplex-3 observed mass: 4774.6 m/z (theoretical mass: 4774.2 m/z), **2**/Seq-6 adduct observed mass: 2951.5 m/z (theoretical mass: 2951.21 m/z).

Lastly, HPLC analysis of Duplex-4 (**Figure 5.16**) gave two peaks at RT 7.07 min (Seq-8) and RT 7.41 min (Seq-7) (**Figure 5.20A**) confirmed by MALDI-TOF-MS as Seq-7 with an observed mass of 2378.7 m/z (theoretical mass: 2379.6 m/z) and Seq-8, 2394.9 m/z (theoretical mass: 2394.6 m/z) (**Figure 5.20C**). After incubation with **2** at 25 °C in a 4:1 molar ratio (**2**/DNA), a new minor peak at RT 18.30 min was observed although reaction did not go to completion with ~7% adduct formation (**Figure 5.20B**). The stoichiometry of the peak at RT 18.30 min was confirmed by MALDI-TOF-MS analysis as the 1:1 **2**/Seq-8 adduct with a detected mass of 2951.7 m/z (theoretical mass: 2951.21 m/z) (**Figure 5.20C**).

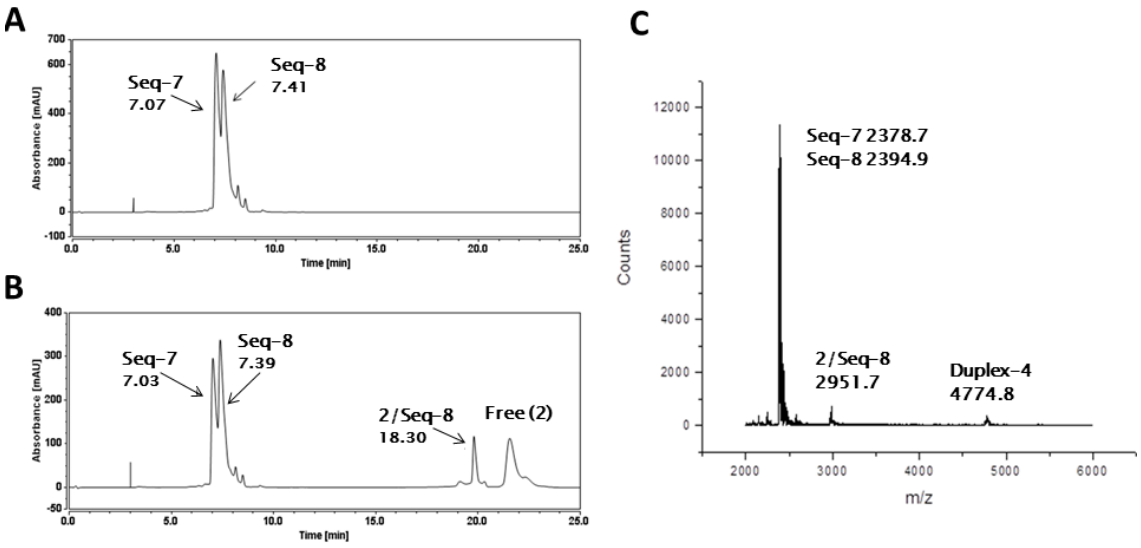


Figure 5.20: HPLC chromatograms; **A**, Annealed Duplex-4 at RT 7.07 min (Seq-8) and RT 7.4 min (Seq-7); **B**, Annealed Duplex-4 after incubating with **2** for 24 hours showing a new adduct peak at RT 18.30 min (**2**/Seq-8), with the reaction not complete after 24 hours; **C**, MALDI-TOF spectrum of **2** with Duplex-4, confirming the stoichiometry of the adduct formed; single strands of Duplex-4, observed masses: Seq-7, 2378.7 m/z and Seq-8, 2394.9 m/z (theoretical mass: Seq-7 2379.6 m/z and Seq-8 2394.6 m/z), and Duplex-4 observed mass: 4774.8 m/z (theoretical mass: 4774.2 m/z), **2**/Seq-8 adduct observed mass: 2951.7 m/z (theoretical mass: 2951.21 m/z).

5.7.2 Summary of the interaction of SJG-136 (2) with double-stranded inosine-modified AP-1 sequences

In summary, the results obtained from the interaction of **2** with Duplexes 1-4 (**Figure 5.16**) were consistent with those obtained for AP-1 Hairpins 1-4 (**Figure 5.8**). Duplex-1 and Duplex-3 both contain a guanine at the 5'-terminus and **2** was able to react with both.

5.8 Interaction of anthramycin (1) and GWL-78 (3) with double-stranded inosine modified Duplex-1 sequence

Based on the data obtained for AP-1 Hairpin-1, anthramycin (**1**) was not expected to react with Duplex-1 as it lacks a substituent at the C8-position that would help stabilise its accommodation in the DNA minor groove. In contrast, **3** was expected to react due to its C8-*bis*-pyrrole substituent.

5.8.1 Results and discussion

The working solution of inosine modified Duplex-1 was prepared according to the procedure described in Section 7.2.1.5. Anthramycin and GWL-78 working solutions were prepared as described in Sections 7.2.1.6 and 7.2.1.7. PBD/DNA complexes were prepared according to the protocol described in Section 7.2.1.10. Briefly, a PBD working solution of 100 µM was added to a DNA working solution of 25 µM in a 4:1 ratio (PBD/DNA). The mixture was mixed and incubated for 24 hours at 25 °C. Following the incubation, the samples were subjected to ion pair RP-HPLC/MALDI-TOF-MS analysis as described in Sections 7.2.2 and 7.2.4 in detail.

The results shown in **Figure 5.21A-C** confirm that there is no reaction of **1** with Duplex-1 after 24 hours incubation. However, **Figures 5.21D-F** show that incubation of **3** with Duplex-1 for 24 hours provides an additional peak in the chromatogram at RT 14.27 min (**Figure 5.21E**), although the reaction was incomplete with ~8% adduct formation. The stoichiometry of the adduct formed was confirmed to be the 1:1 **3**/Seq-1 adduct by MALDI-TOF-MS analysis with a detected mass of 2985.7 m/z (theoretical mass 2985.2 m/z) (**Figure 5.21F**).

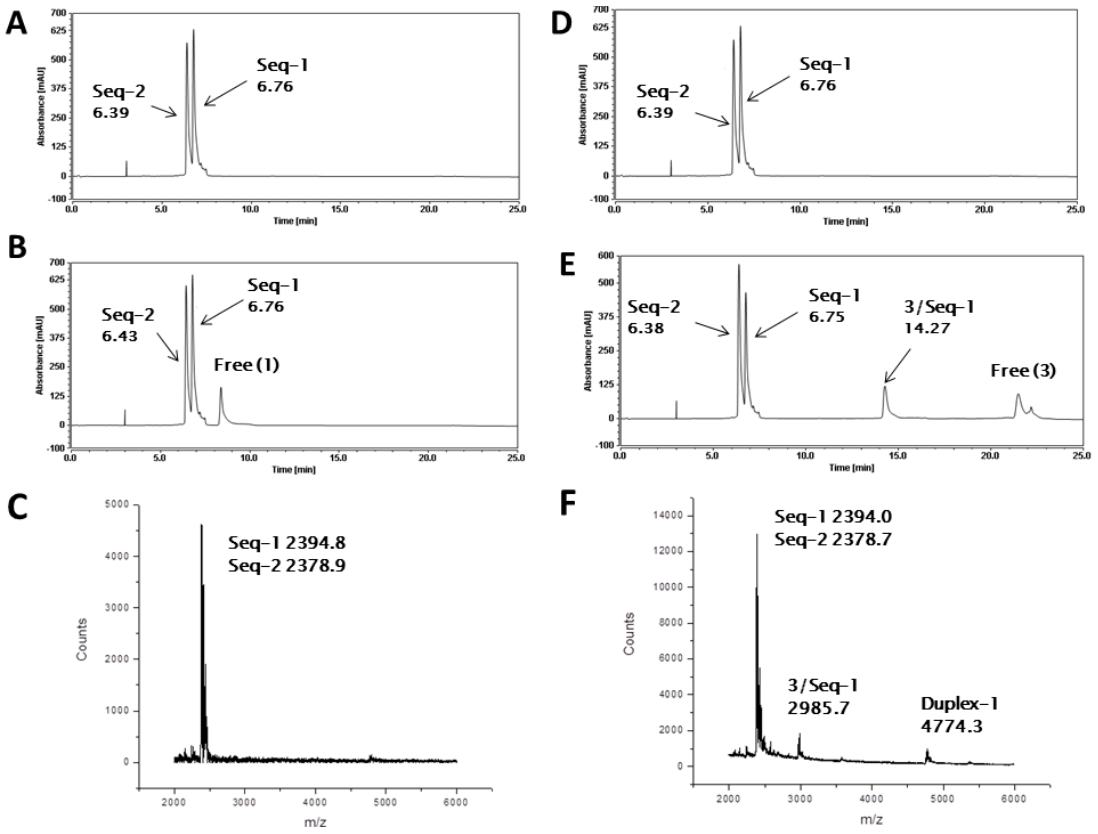


Figure 5.21: HPLC chromatograms: **A**, Annealed Duplex-1 at RT 6.39 min (Seq-2) and RT 6.76 min (Seq-1); **B**, After incubation with **1** for 24 hours showing no reaction; **C**, MALDI-TOF spectrum of **1** with Duplex-1 confirming no reaction; Single strands of Duplex-1 observed masses: Seq-2, 2378.9 m/z and Seq-1, 2394.8 m/z (theoretical masses: Seq-2, 2379.6 m/z and Seq-1, 2394.6 m/z); **D**, Annealed Duplex-1 at RT 6.39 min (Seq-2) and 6.76 min (Seq-1); **E**, Annealed Duplex-1 after incubating with **3** for 24 hours showing a new peak at RT 14.27 min, although reaction was incomplete (~8% adduct formed); **F**, MALDI-TOF spectrum of **3** with Duplex-1 confirming the stoichiometry as the 1:1 **3**/Seq-1 adduct; Single strands of Duplex-1 observed masses: Seq-2, 2378.7 m/z and Seq-1, 2394.0 m/z (theoretical masses: Seq-2, 2379.6 m/z and Seq-1, 2394.6 m/z) and double-stranded Duplex-1 observed mass of 4774.3 m/z (theoretical mass: 4774.2 m/z), **3**/Seq-1 adduct observed mass of 2985.7 m/z (theoretical mass: 2985.2 m/z).

5.8.2 Summary of the interaction of anthramycin and GWL-78 with double-stranded inosine modified Duplex-1 sequence

Taken together, these results suggest that the C8-substitution status of a PBD molecule can have a significant effect on its binding behaviour to a target DNA sequence. For

example, **1** lacks a C8-substituent and is unable to interact with the guanine located at the 5'-terminus of AP-1 Hairpin-2 and Duplex-1, whereas **3** can bond to both sequences due to its C8-*bis*-pyrrole substituent. For similar reasons, SJG-136 (**2**) can also bond to the 5'-terminal guanine of AP-1 Hairpin-2 and Duplex-1 because the second PBD unit joined through the C8-O-$(CH_2)_3$-O-C8' linker can also provide stabilisation by interacting in the minor groove, even if there is no suitable guanine available for alkylation by the second PBD unit.

5.9 Molecular dynamics simulations*

Molecular dynamics simulations were carried out to attempt to predict the most preferred reacting guanine(s) in the AP-1 hairpin sequence and correlate with biophysical results. In these simulations, **2** was covalently bound to every potentially reacting guanine base (*i.e.*, G1, G7, G12 and G17) in an effort to rationalise the structures of the adducts formed in the HPLC studies. As **2** is known to form mono-adducts and inter- and intrastrand cross-links[191b, 205], all of these adduct types were investigated and, in the case of mono-adducts, both loop-facing (*i.e.*, forward) and non-loop-facing (*i.e.*, reverse) orientations were considered. The study design also considered potential reaction sites based on the span of the molecule. For example, in the case of G7 and G12, mono-adducts with the bulk of the molecule pointing toward the loop were not analysed due to potential steric hindrance with the TTT-loop. However, adducts at G17 in both orientations were investigated despite the fact that the significant bulk of a 3'-oriented adduct would be positioned outside of the minor groove environment.

Potential energy calculations (kcal/mol) on each adduct were undertaken to establish the most preferred reacting guanines for either mono-adduct or cross-link formation. Although **2** is known to prefer to form cross-linked adducts over a six or seven base-pair sequence (*i.e.*, Pu-GATC-Py, Pu-GATG-Py, Pu-GAATG-Py or Pu-GAATC-Py[205]), initial potential energy calculations suggested that the shorter interstrand cross-linked adduct spanning 5'-GAC-3' (*i.e.*, G1-G17) is preferred by 12.41 kcal/mol compared to

** Molecular dynamics simulations studies were undertaken by Dr Paul Jackson*

the extended G7-G17 interstrand cross-linked adduct (*i.e.*, 5'-CATTG-3') (**Figure 5.2**), despite the G1-bonded PBD unit interacting with two rather than the usual three base pairs at the terminal 5'-guanine. Although this result was surprising, especially as the extended 5'-ATT-3' bases in the centre of the hairpin should form van der Waals interactions with the C8/C8'-linker of **2**, it is possible that DNA breathing and base-pair separation affects the ability of **2** to stretch across the extended 5'-CATTG-3' sequence which may have a detrimental effect on potential energy values. Furthermore, previous studies have suggested that the A-ring-3' orientation for a PBD (*i.e.*, 5'-GAC-3') is preferred to A-ring-5' (*i.e.*, 5'-CATTG-3'), and although less favourable than Pu-GATC-Py, a 5'-GAC-3' adduct of **2** has been previously reported[205]. For possible mono-adducts, the potential energy calculations suggested that G7 is the preferred reacting guanine (**Figure 5.22**). During simulations of this adduct, a stabilising hydrogen bond was observed to form between the N10-H of the PBD and the N3 of A14. Furthermore, the second PBD unit of the dimer formed favourable non-covalent interactions with the central AT-tract of the sequence (*i.e.*, 5'-ATT-3'), with both interactions likely to contribute to enhanced stability within the minor groove environment. The potential energy calculations suggested a hierarchy of mono-adduct formation of G7 > G17 (in the *forward* orientation) > G1 > G12 > G17 (in the *reverse* orientation). This ranking was broadly consistent with the observations from the inosine replacement experiments which suggested that G7 and G17 are favoured for nucleophilic attack based on the faster reaction rates of the parent AP-1 sequence, AP-1 1, 12I and AP-1 7, 12I, AP-1 Hairpins 2 and 4, and Duplexes 2 and 4. Conversely, potential energy calculations predicted that less favourable binding sites would include G12 where the TTT-loop is likely to inhibit interaction of **2**, or G17 (in the *reverse* orientation) where the second PBD unit of the dimer would be positioned outside of the groove where it would not be able to form non-covalent interactions with DNA bases in the minor groove.

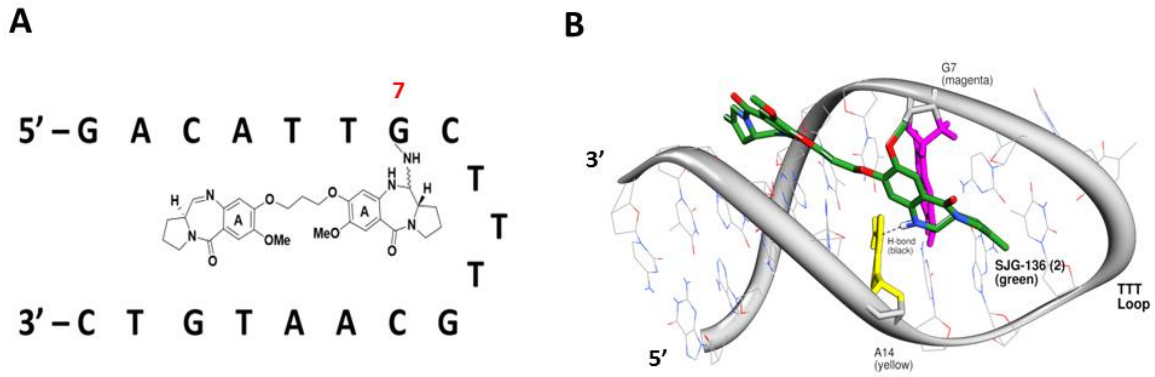


Figure 5.22: Low energy snapshot of a molecular dynamics simulation showing the accommodation of **2** (green sticks) in the minor groove while covalently bonded to G7 of the parent AP-1 sequence (**Figure 5.2**). The NH group of the PBD formed a hydrogen bond with the adjacent A14 adenine base (yellow), and the molecule maintained good isohelicity with the minor groove for the duration of the simulation (10 ns).

It is noteworthy that potential energy calculations for the bonding of a PBD molecule to G1, both in mono-adduct and cross-linking modes, suggest that it is viable as a reacting guanine. G1 adducts were observed by HPLC for the reaction of both **2** (**Figure 5.23**) and **3** (**Figure 5.24C** and **D**) with 5'-GACATTIC-TTT-IAATITC-3' (AP-1 Hairpin-1) where G1 is the only available guanine for reaction, and also in the sequence 5'-GATATTAT-TTT-ATAATATC-3' (AP-1 1G hairpin) which is also devoid of other reactive guanines. However, when anthramycin (**1**) was reacted with the same sequences (**Figure 5.24A** and **B**), no adducts were formed. Although molecular dynamics simulations of **1** covalently bound to G1 suggested that an adduct in the C11-*S* configuration is theoretically possible, it is likely that the lack of non-covalent interactions from a C8-substituent, as available in case of **2** and **3**, leads to low overall stabilisation of the adduct. Further evidence for this hypothesis was obtained through molecular dynamics simulations of **2** and **3** bound non-covalently to AP-1 Hairpin-1, where non-covalent interactions from the C8-pyrrole chain (in the case of **3**) and the non-alkylating PBD unit (in the case of **2**), oriented the alkylating PBD component over G1 for a covalent attack. However, in the case of **1** which does not have a C8-side chain, it moved up and down the minor groove during the simulation and failed to align over G1 for attack by the C2-amino group.

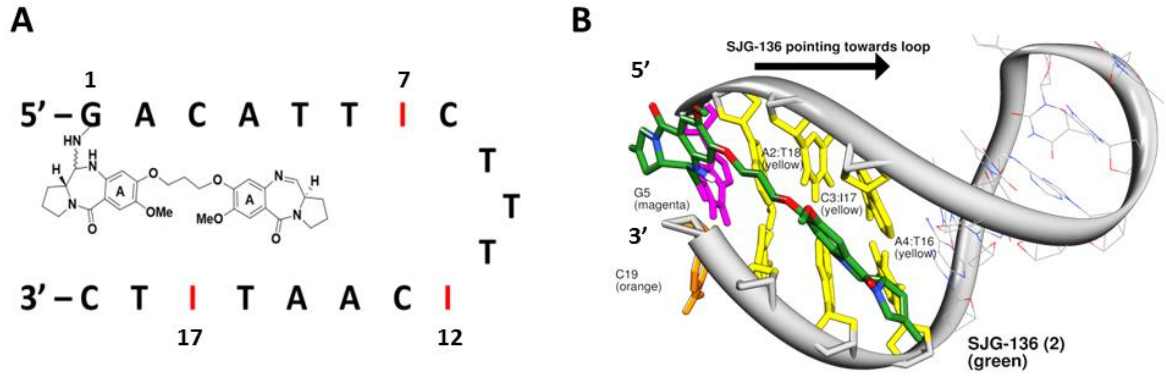


Figure 5.23: **A**, Schematic model of **2** bound covalently to G1 of AP-1 Hairpin-1 with the bulk of the molecule positioned within the minor groove and pointing towards the TTT-loop (*i.e.*, the forward direction); **B**, Low energy snapshot of a molecular dynamics simulation (10 ns) of the adduct shown in **Panel A**, illustrating the good accommodation of **2** (green sticks) in the minor groove and forming non-covalent interactions between the central methylene linker of the ligand and the A2:T18 and C3:I17 base-pairs. **2** is covalently bound to the 5'-terminal G1 of AP-1 Hairpin-1 (**Figure 5.8**).

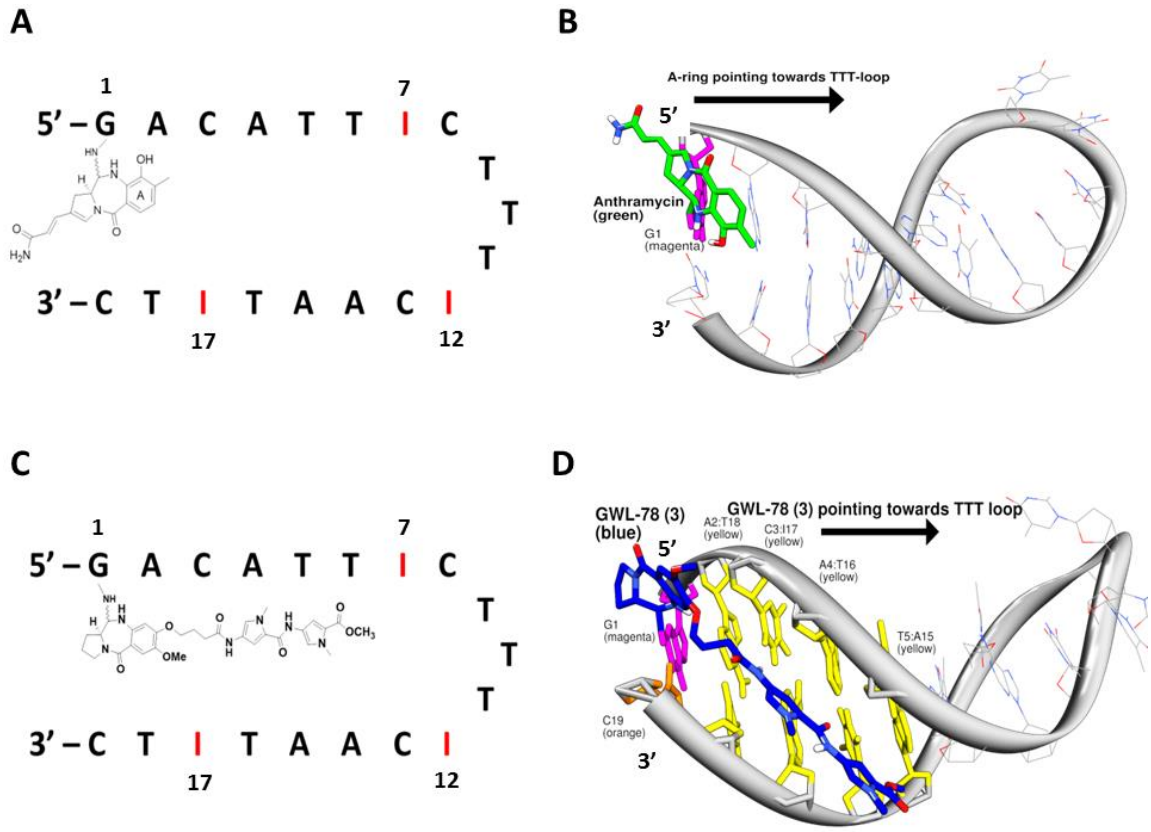


Figure 5.24: **A,** Schematic model of **1** (green) covalently bound to G1 (magenta) of AP-1 Hairpin-1 (**Figure 5.8**) with the A-ring pointing towards the TTT-loop (A-ring-3' orientation); **B,** Low energy snapshot of 10 ns molecular dynamics simulation illustrating the C2-tail of **1** orienting outside of the DNA, suggesting a DNA triplet is necessary for accommodation of the molecule; **C**, Schematic model of **3** covalently bound to G1 of AP-1 Hairpin-1 with the A-ring pointing towards the TTT-loop; **D**, Low energy snapshot of 10 ns molecular dynamics simulation of **3** (blue) covalently bound to G1 (magenta) of AP-1 Hairpin-1, illustrating the comfortable accommodation of the poly-pyrrole tail of **3** in the minor groove. It is likely that the poly-pyrrole tail forms non-covalent interactions with 5'-ACAT-3' (yellow), guiding the PBD core to G1.

Simulations of **1**, **2** and **3** covalently bound to G1 of AP-1 Hairpin-1 were also undertaken with the ligands pointing away from the loop (**Figures 5.25**, **5.26** and **5.27**) to investigate the orientation of the G1 adduct. Potential energy calculations of **2** and **3** in both orientations suggest a potential energy difference of between 40 kcal/mol and 43 kcal/mol, respectively, in favour of the loop-facing orientation. This most likely reflects the unfavourable energetics when the bulk of the molecule is pointing out of the DNA helix into a water environment rather than laying in the minor groove where it can provide adduct stabilisation through non-bonding interactions. As anticipated, the energy

difference between the two possible orientations of **1** covalently bonded to G1 was less (*i.e.*, -2736.10 kcal/mol *versus* -2733.60 kcal/mol) due to the lack of a C8-substituent that can provide adduct stabilisation through non-covalent interactions.

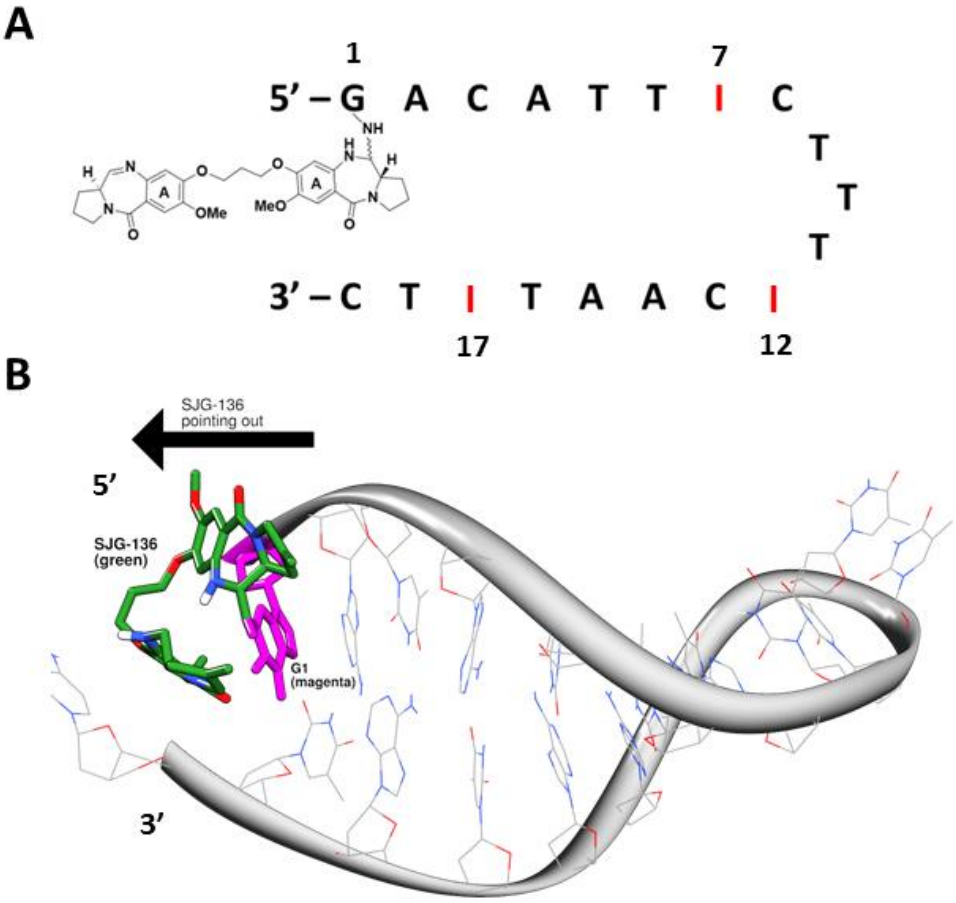


Figure 5.25: **A,** Schematic model of **2** (green) covalently bound to G1 (magenta) of AP-1 Hairpin-1 with the A-ring pointing away from the TTT-loop (A-ring-3' orientation) and the second PBD of the dimer orienting outside of the minor groove; **B**, Low energy snapshot of 10 ns molecular dynamics simulation illustrating the second PBD of **2** orienting outside of the DNA minor groove.

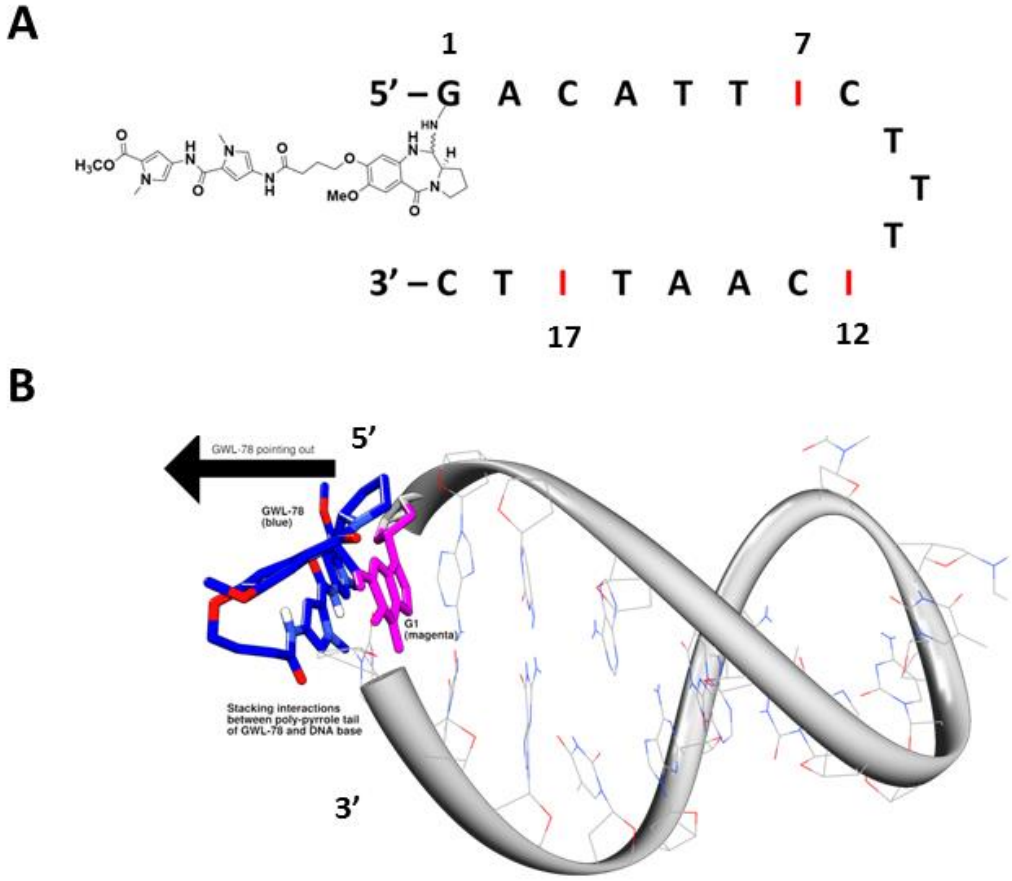


Figure 5.26: A, Schematic model of **3** (blue) covalently bound to G1 (magenta) of AP-1 Hairpin-1 with the A-ring pointing away from the TTT-loop (A-ring-3' orientation) and poly-pyrrole chain orienting outside of the minor groove; **B**, Low energy snapshot of 10 ns molecular dynamics simulation illustrating the poly-pyrrole chain of **3** orienting outside of the DNA minor groove and forming stacking interactions with the G1 base.

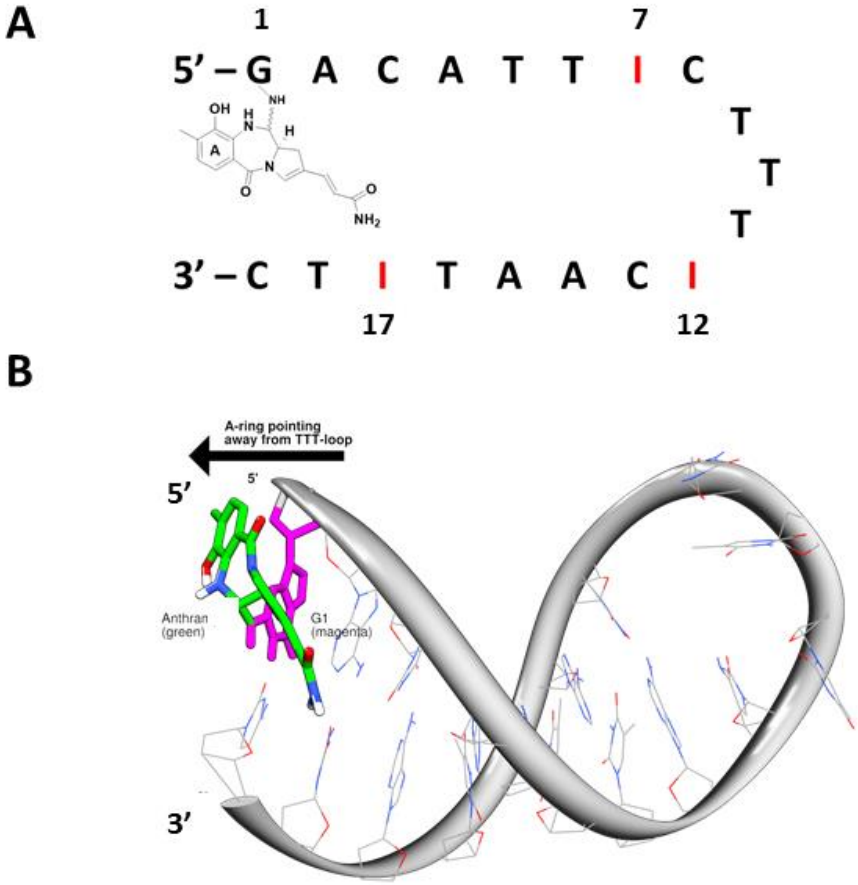


Figure 5.27: A, Schematic model of **1** (light green) covalently bound to G1 (magenta) of AP-1 Hairpin-1 with the A-ring pointing away from the TTT-loop (A-ring-3' orientation) **B**, Low energy snapshot of 10 ns molecular dynamics simulation illustrating the PBD (**1**) bound to the G1 base through interactions with a doublet rather than a traditional triplet.

5.10 Conclusion

In summary, the data obtained from the study on the hairpins outlined in **Figure 5.3**, suggest that the G7-G17 interstrand cross-link may be the preferred formed adduct between SJG-136 and the parent AP-1 hairpin sequence as only for the AP-1 1, 12I sequence the same extent of reaction was observed as for the parent AP-1 sequence during the HPLC/MS study.

Moreover, this is the first report of PBD molecules covalently binding to a guanine base located at the end of a DNA hairpin or duplex sequence. To date, PBDs have been reported to span three DNA base pairs with a preference for the 5'-Pu-G-Pu-3' motif, and were not thought to be able to bond to a reactive guanine at a 3'- or 5'-terminus. The results presented in this study clearly demonstrate the ability of the C8-conjugated PBD GWL-78 (**3**) and the PBD dimer SJG-136 (**2**) to covalently interact with a terminal guanine, whereas the monomeric anthramycin (**1**) does not react. This is most likely due to the C8-substituent of **3**, or one of the C8-linked PBD units of **2**, interacting in the minor groove to provide stability for the adduct. This has possible implications for the mechanism of the cellular cytotoxicity and *in vivo* anti-tumour activity of C8-substituted PBD monomers and C8-linked PBD dimers, as they may be able to covalently interact with guanine bases at the ends of DNA double-strand breaks in cells.

Chapter 6: Overall conclusion and future work

An ion pair reversed phase HPLC and MALDI-TOF-MS assay has been developed to study the interaction of the pyrrolobenzodiazepine (PBD) dimer SJG-136 with the cognate target DNA sequences of the oncogenic transcription factors NF-κB, EGR-1, AP-1 and STAT3. Initially, the assay was validated by using it to investigate the interaction of the PBD monomers GWL-78 and KMR-28-39 with the target sequences. These molecules had already been reported to interact with the consensus sequences of the transcription factors NF-Y and NF-κB[90, 102], respectively. The data obtained (see Chapter 2) indicated that the analytical method developed may be capable of identifying and quantitating the rate and extent of reaction of these PBD monomers with NF-κB, EGR-1, AP-1 and STAT3 cognate sequences. After validating the HPLC/MS assay, it was then used to evaluate the ability of SJG-136 to bind to these transcription factor sequences with a view to using the information obtained to help interpret the activity of SJG-136 in various tumour cell lines. For example, SJG-136 is particularly potent in NF-κB-dependent (*e.g.*, leukaemia and lung cancer) and AP-1-dependent cell lines (*e.g.*, colon carcinoma) according to NCI 60 cell line screen data[158], while the STAT3-dependent breast cancer cell line MDA-MB-435 appeared to be one of the least sensitive cell lines toward SJG-136.

The data presented in Chapter 3 represent the first report in which SJG-136 appears to bind to the consensus sequences of NF-κB (two possible consensus sequences were studied), EGR-1, AP-1 and STAT3 based on the HPLC/MS methodology developed. Significant differences in the extent and rate of reaction with these individual sequences were observed. For example, SJG-136 reacted rapidly with the NF-κB-2 and AP-1 sequences, but more slowly with the EGR-1 sequence. This observation however may be due to the insufficient peak resolution of the applied method. Future studies will include improvement of the HPLC method in order to address this issue. The reaction rate was slowest with the NF-κB-1 and STAT3 sequences. In accord with previous reports that SJG-136 can form intra- and interstrand cross-links and mono-alkylated adducts within the DNA minor groove[96, 136], multiple adducts were observed for the reaction of SJG-136 with some of the transcription factor sequences due to their structures which contain multiple GC base pairs. For example, SJG-136 formed three distinct adducts with the

STAT3 consensus sequence and two with the NF-κB-1 sequence, while only a single adduct was observed to form with the NF-κB-2, EGR-1 and AP-1 sequences. These results may be useful in future studies to further understanding of the mechanism of action of SJG-136, and to help correctly interpret the results of biochemical and pharmacological assays.

SJG-136 was also evaluated for its ability to affect protein expression in the human colon carcinoma cell line HT-29 (AP-1 dependent) and the human breast cancer cell line MDA-MB-231 (STAT3 dependent) using the polymerase chain reaction (PCR) and Western blotting methodologies, and the results are presented in Chapter 4. Although SJG-136 was observed to significantly down-regulate AP-1- and STAT3-dependent protein expression (*e.g.*, *Bcl-2*, *fascin*, *NNMT*, *VEGF*, *CREB5* and *p53*) in the PCR study, equivalent down-regulation was not observed by Western blotting, and only cyclin D1 protein expression levels were significantly reduced. This discrepancy was surprising, and future studies could include time-course studies using a wide range of inducer agents applied to the cells for different time intervals as well as at different concentrations. Furthermore, time-course experiments, and especially longer incubation times adjusted to the half-lives of the proteins of interest, could be carried out to provide further information. Also, a broader range of proteins implicated in tumour development and progression could be investigated.

The surprising observations made with the AP-1 cognate sequence (*i.e.*, fastest reaction with SJG-136 with 100% adduct formation in the HPLC/MS study) led to a DNA binding site analysis on this particular sequence in order to identify the reactive guanines responsible for the observed rate and extent of reaction. For this reason, the study further included inosine replacement experiments as a strategy for trying to elucidate the type of adducts formed within the AP-1 sequence. The guanines were consecutively replaced with inosine bases which lack the nucleophilic C2-NH_2 functionalities and prevent covalent interaction of SJG-136. The results presented in Chapter 5 suggest that the interstrand cross-link at G7-G17 may be the preferred formed adduct between SJG-136 and the parent AP-1 hairpin sequence. Moreover, the results presented in Chapter 5 demonstrate for the first time, that PBD monomers and dimers can covalently bond to a guanine base located at the end of DNA hairpin or duplex fragments. Before these results

were obtained, PBDs were thought to require three DNA bases (with a preference for a 5'-Pu-G-Pu-3' motif) for covalent attachment within the DNA minor groove[298]. Furthermore, for C8/C8'-linked dimers such as SJG-136 and DSB-120, 6 or 7 base pairs were reported to be essential for covalent binding to DNA[97]. However, the discovery that PBD monomers and dimers can covalently bond to terminal guanine residues highlights the importance of the C8-substituent in dictating the formation of particular types of adduct. The biological relevance of these findings may have implications for the mechanism of the cellular cytotoxicity and *in vivo* anti-tumour activity of C8-substituted PBD monomers and C8-linked PBD dimers. To date, PBDs have been reported to exert their anti-tumour activity through the formation of mono-alkylated or intra- or interstrand cross-linked adducts within stretches of DNA, thus inhibiting crucial processes including transcription factor binding, enzyme processing, transcription, or arrest of the replication fork. However, the data presented in Chapter 5 suggest that a mechanism involving binding of C8-substituted PBDs to the ends of DNA double-strand breaks in cells could also occur. This finding may have significance for the PBD-based Antibody-Drug Conjugates (ADCs) currently in clinical evaluation, and numerous others at the pre-clinical stage.

Other future studies could include evaluation of SJG-136 in NF-κB and EGR-1 over-expressing cell lines such as leukaemia and ovarian tumour cells in order to correlate protein expression with the observations made during the HPLC/MS studies. Furthermore, the HPLC/MS assay developed during this project could be applied to novel PBD monomers and dimers produced in the future to investigate their ability to inhibit the transcription of specific genes.

Chapter 7: Experimental (biophysical and biological evaluation)

7.1 Consumables and reagents used for biophysical and biological experiments

Consumables and reagents including their manufacturer used for biophysical and biological experiments are listed in **Table 7.1**.

Table 7.1: Consumables and reagents including their manufacturer used for biophysical and biological experiments.

Manufacturer	Consumables/reagents
Sigma-Aldrich, Dorset, UK	Tris-HCl, ethylenediamine-tetraacetic acid (EDTA), ammonium acetate, dimethylsulfoxide (DMSO), 1 M triethylammonium bicarbonate (TEAB) buffer, 3-hydroxypicolinic acid (HPA), 1 M triethylammonium acetate buffer (TEAA), ammonium citrate dibasic (DAC), potassium hydroxide (KOH), cacodylic acid, trypan blue solution, agarose, trizima HCl, Laemmli sample buffer, goat-anti-rabbit IgG secondary antibody, Carestream Kodak autoradiobiography GBX developer/replenisher, Carestream Kodak autoradiobiography GBX fixer
Fisher Scientific, Loughborough, UK	Sodium chloride (NaCl), biology grade nuclease-free water, acetonitrile, water HPLC grade, potassium chloride (KCl), tris-base, methanol, sodium dodecylsulfate (SDS), glycerine, bovine serum albumin (BSA)

Millipore, Millipore Corporation, Billerica, MA, USA	C_{18} ZipTips™
Invitrogen, Paisley, UK	Ambion DEPC (diethylpolycarbonate) - treated water, DMEM Glutamax medium (Dulbecco's Modified Eagle Medium containing 4.5 g/L D-Glucose and 110.0 mg/L Sodium Pyruvate), fetal bovine serum (FBS), minimum essential medium non-essential amino acids (MEM NEAA 100 X), Penstrep (10,000 Units/mL Penicillin, 10,000 µg/mL streptomycin, Dulbecco's Phosphate Buffered Saline (PBS [-] $CaCl_2$, [-] $MgCl_2$), 0.25% trypsin-EDTA, 10 X PCR reaction buffer, 10 mM dNTP mix, 50 mM $MgCl_2$, Platinum Taq polymerase, 10 X Tris-acetate-EDTA (TAE) buffer, 100 bp ladder, 4-12% NuPage Novex Bis-Tris SDS-PAGE gel, NuPage MOPS SDS running buffer 20 X, Novex Sharp Pre-stained protein standard
American Type Culture Collection (ATCC), Manassas, VA, USA	MDA-MB-231, HT-29, Mc Coy's 5A medium
Quiagen, Hilden, Germany	RNeasy® Plus Mini kit
Applied Biosystems, California, USA	High Capacity RNA-to-cDNA kit
Alpha Laboratories, Eastleigh, UK	0.2 mL thin-walled, dome-capped polymerase chain reaction (PCR) tubes
Roche Diagnostics, Mannheim, Germany	PCR grade water, Fast Start Universal Probe Master Mix
New England Biolabs, Ipswich, USA	6 X blue gel loading buffer
Biotinum, California, USA	GelRed
VWR, Leicestershire, UK	Triton X-100
Bio-Rad, Hemel Hempstead, UK	Bio-Rad DC Protein assay reagent, Quick

	Draw Filter (QDF) paper, StrepTactin-HRP conjugate
Thermo Scientific, Loughborough, UK	Restore Western blot stripping buffer
GE Healthcare Life Sciences, Buckinghamshire, UK	Hybond-ECL nitrocellulose membrane, ECL detection kit, Amersham Hyperfilm ECL
Cell Signalling Technology, Cambridge, UK	Rabbit monoclonal anti-cyclin D1,-Bcl-2
Abcam, Cambridge, UK	Rabbit monoclonal anti-pSTAT1,-pSTAT3,-tSTAT1,-tSTAT3,-p53,-VEGF,-α-tubulin

7.2 Experimental (biophysical evaluation of SJG-136)

7.2.1 Sample preparation for HPLC analysis

7.2.1.1 General

The single-stranded oligonucleotides were purchased from Eurogentec, Southampton, UK in lyophilised forms and prepared for experiments as described in Sections 7.2.1.2 – 7.2.1.5. Anthramycin and SJG-136 were supplied by Spirogen Ltd. and prepared as described in Sections 7.2.1.6 and 7.2.1.9. Identity and purity was confirmed by LC-MS analysis prior to further processing. The PBD monomers GWL-78 and KMR-28-39 were synthesised by the Thurston and Rahman group and prepared as described in Sections 7.2.1.5 and 7.2.1.6. Consumables and reagents including their manufacturer used for HPLC experiments are listed in **Table 7.1**. Preparations of solutions required for HPLC analysis are described in **Table 7.2** in detail. All prepared solutions were stored at room temperature.

Table 7.2: Composition and preparation of solutions required for HPLC analysis.

Solution	Composition
Annealing buffer	60.57 mg Tris-HCl (121.14 g/mol) (10 mM), 146.1 mg NaCl (58.44 g/mol) (50 mM), and 14.612 mg EDTA (292.24 g/mol) (1 mM) were weight out, mixed and dissolved in 20 mL of deionised water. The volume was made up to 50 mL with deionised water. pH was adjusted to 8.5.
100 mM ammonium acetate buffer	385.4 mg ammonium acetate (77.08 g/mol) was dissolved in 20 mL deionised water. The volume was made up to 50 mL with deionised water.
20 mM ammonium acetate buffer	77.08 mg ammonium acetate (77.08 g/mol) was dissolved in 20 mL deionised water. The volume was made up to 50 mL with deionised water.
100 mM TEAB buffer	100 mL 1 M TEAB buffer were mixed with 900 mL HPLC grade water.
40% acetonitrile in water	400 mL of 100% acetonitrile were mixed with 600 mL HPLC grade water.

7.2.1.2 Preparation of working solutions of hairpin oligonucleotides

Each single-stranded oligonucleotide was dissolved in annealing buffer and 100 mM ammonium acetate in a 1:1 ratio to form stock solutions of 1 mM. To ensure hairpin formation, the solutions were heated to 85 °C for 10 min in a heating/cooling block (Grant Bio, UK). The solutions were allowed to cool slowly down to room temperature followed by storage at -20 °C overnight to ensure completion of the annealing process. Working solutions of hairpin oligonucleotides of 25 µM were prepared by diluting the stored stock solutions with 100 mM ammonium acetate and storage at -20 °C.

7.2.1.3 Preparation of working solutions of duplex oligonucleotides

The single-stranded oligonucleotides were dissolved in annealing buffer and 100 mM ammonium acetate in a 1:1 ratio to form stock solutions of 1 mM. Double-stranded oligonucleotides were prepared by heating the single-stranded DNA sequences to 85 °C for 10 min in a heating/cooling block (Grant Bio, UK). The solutions were allowed to cool slowly down to room temperature followed by storage at -20 °C overnight to ensure completion of the annealing process. Working solutions of duplex oligonucleotides of 25 µM were prepared by diluting the stored stock solutions with 20 mM ammonium acetate and storage at -20 °C.

7.2.1.4 Preparation of working solutions of inosine modified hairpin oligonucleotides

Inosine modified oligonucleotides have been used for binding site analysis of SJG-136 within the AP-1 consensus sequence. Inosine modified sequences were purchased from Eurogentec, Southampton, UK in lyophilised forms. Preparation of working solutions of inosine modified hairpin oligonucleotides is described in Section 7.2.1.2 in detail.

7.2.1.5 Preparation of working solutions of inosine modified duplex oligonucleotides

Single-stranded inosine modified sequences were purchased from Eurogentec, Southampton, UK in lyophilised forms. Preparation of working solutions of inosine modified duplex oligonucleotides is described in Section 7.2.1.3 in detail.

7.2.1.6 Preparation of working solution for PBD monomer anthramycin

Anthramycin was dissolved in DMSO to form stock solution of 10 mM which was stored at -20 °C. Ligand working solution of 100 µM was prepared by diluting the stock solution with nuclease-free water. The final working solution was stored at -20 °C and thawed to room temperature when required.

7.2.1.7 Preparation of working solution for PBD monomer GWL-78

GWL-78 was dissolved in DMSO to form stock solution of 10 mM which was stored at -20 °C. Ligand working solution of 100 µM was prepared by diluting the stock solution with nuclease-free water. The final working solution was stored at -20 °C and thawed to room temperature when required.

7.2.1.8 Preparation of working solution for PBD monomer KMR-28-39

The PBD monomer KMR-28-39 was dissolved in DMSO to form the stock solution of 10 mM which was stored at -20 °C. Ligand working solution of 100 µM was prepared by diluting the stock solution with nuclease-free water. The final working solution was stored at -20 °C and thawed to room temperature when required.

7.2.1.9 Preparation of working solution for PBD dimer SJG-136

The PBD dimer SJG-136 was dissolved in DMSO to form stock solution of 10 mM which was stored at -20 °C. Ligand working solution of 100 µM was prepared by diluting the stock solution with nuclease-free water. The final working solution was stored at -20 °C and thawed to room temperature when required.

7.2.1.10 Preparation of ligand/DNA complexes

Ligand/DNA complexes were prepared in Eppendorf tubes by adding a PBD working solution of 100 µM to a hairpin or duplex oligonucleotide working solution of 25 µM in a 4:1 (PBD/DNA) ratio. The mixture was agitated for 5-10 s using a vortex mixer and incubated in a heating/cooling block for various time intervals at 25 °C and subjected to ion pair reversed-phase HPLC and mass spectrometry analysis as described below.

7.2.2 Ion pair reversed-phase HPLC

Liquid chromatography was performed on a Dionex UltiMate 3000 system (Thermo Scientific, Loughborough, UK) equipped with a 2.1 x 50 mm XBridge™ OST C_{18} column packed with 2.5 µm particles (Waters Ltd., UK). The gradient system used for LC

analysis consisted of 100 mM TEAB as buffer A and 40% acetonitrile in water as buffer B. Flow gradients used for HPLC analysis are described in the individual chapters in detail. Sampler temperature was held at 15 °C and the column temperature at 20 °C. Samples were injected using a full loop injection of 50 µL. UV absorbance was monitored at 254 nm. Sample peaks were integrated and results expressed as "Area Under the Curve" (AUC) using a Chromeleon 7 software Version 7.1.1.1127 (Thermo Scientific, Loughborough, UK).

7.2.3 Sample preparation for mass spectrometry (MALDI-TOF-MS) analysis

7.2.3.1 General

Consumables and reagents including their manufacturers used for mass spectrometry experiments are listed in **Table 7.1**. Preparation of solutions required for mass spectrometry analysis is described in **Table 7.3** in detail. All prepared solutions were stored at room temperature.

Table 7.3: Composition and preparation of solutions required for mass spectrometry analysis.

Solution	Composition
0.1 M TEAA buffer	10 mL 1 M TEAA was mixed with 90 mL HPLC grade water.
50% acetonitrile in water	50 mL 100% acetonitrile was mixed with 50 mL HPLC grade water.
50% acetonitrile 0.1M TEAA buffer in a 1:1 ratio	0.1 M TEAA and 50% acetonitrile were mixed in a 1:1 ratio.
MALDI matrix	9 parts 50 mg/mL HPA in 50% acetonitrile in water was mixed with 1 part 50 mg/mL DAC in HPLC grade water. 0.1 part formic acid was added to the solution. Matrix was prepared freshly prior to MALDI-TOF-MS analysis.

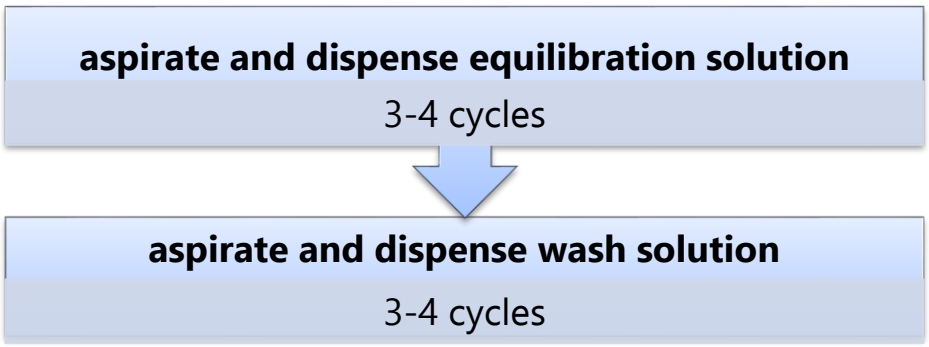
7.2.3.2 Sample preparation for MALDI-TOF analysis

Anthramycin, GWL-78, KMR-28-39, and SJG-136 were prepared and incubated with hairpin or duplex oligonucleotides as described in Section 7.2.1. Incubation was carried out for 24 hours to ensure maximum drug/DNA interaction. The covalently bound drug/DNA adducts samples for MALDI-TOF-MS were diluted with 0.1 M TEAA in 1:1, 1:4, and 1:10 ratios. A ZipTip$_{C18}$™ reversed phase sample preparation methodology was used in order to desalt and concentrate the oligonucleotide samples prior to MALDI-TOF analysis as this procedure provides better data quality (**Figure 7.1**).

Initially, the ZipTip$_{C18}$™ was equilibrated for sample binding by aspirating and dispensing the equilibration solution (which was composed of 50% acetonitrile and 0.1 M TEAA in a 1:1 ratio) followed by washing with wash solution of 0.1 M TEAA. Equilibration and wash steps were repeated 3-4 times. Binding and washing of oligonucleotides was carried out by aspirating the same ZipTip$_{C18}$™ into the sample solution. For maximum sample binding this step was repeated 10 times. Afterwards, the ZipTip$_{C18}$™ was washed for 3 cycles by aspirating and dispensing the wash solution (0.1 M TEAA) followed by a wash with HPLC grade water for 3 cycles. Drug/DNA samples were eluted from the ZipTip$_{C18}$™ and mixed with 0.5 µL matrix. 0.2 µL of sample/matrix mix were spotted onto the MALDI-TOF target plate (MTP AnchorChip 384TF, Bruker Daltonik, Bremen, Germany) and allowed to air-dry prior to MALDI-TOF-MS analysis.

7.2.3.3 MALDI-TOF sample preparation flowchart using ZipTip$_{C18}$™ methodology

ZipTip$_{C18}$™ equilibration for sample binding



Oligonucleotides binding and washing

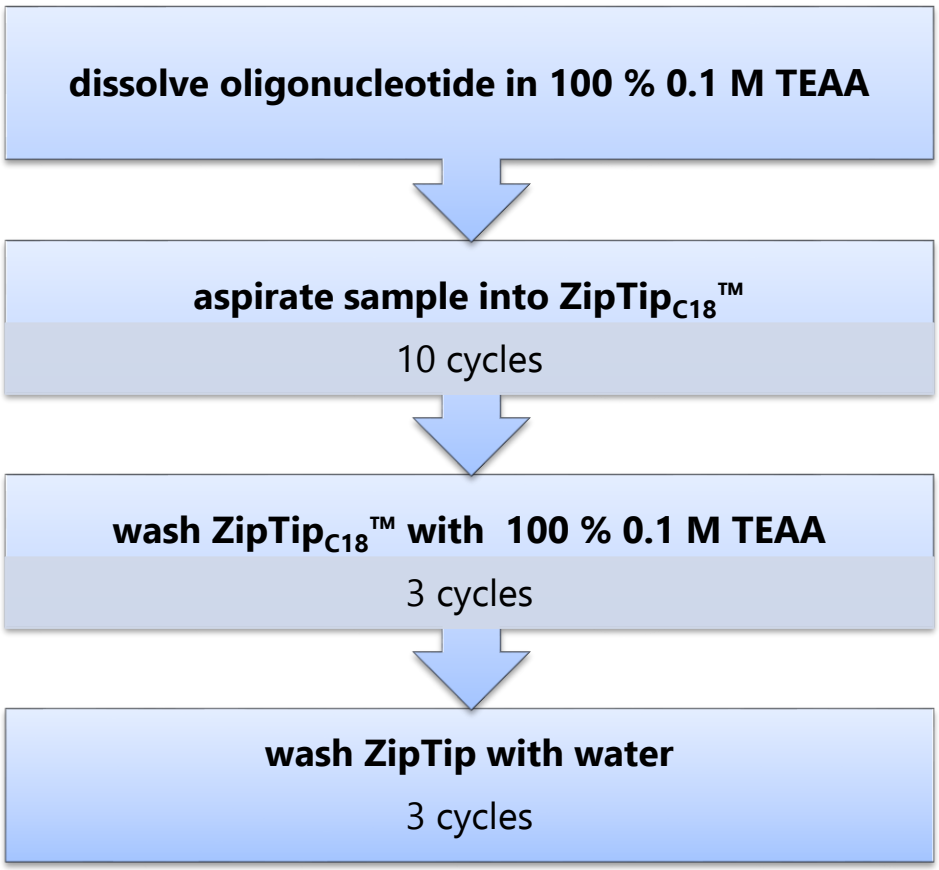


Figure 7.1: Flow chart of sample preparation of oligonucleotides prior to MALDI-TOF-MS analysis using ZipTip$_{C18}$™ methodology.

7.2.4 MALDI-TOF-MS analysis

MALDI-TOF analysis was carried out on a Bruker Daltonics Autoflex™ (Bruker Daltonik, Bremen, Germany) automated high-throughput Matrix-Assisted Laser Desorption/Ionisation Time of Flight (MALDI-TOF) system with a nitrogen laser in positive linear mode using delayed extraction of 500 ns and an accelerating voltage of 25000 V. Acquisition was between 1000-10000 Dalton with 100 shots per spectrum. The instrument was calibrated prior to sample analysis using insulin as the standard. 1 µL insulin was spotted onto the MALDI-TOF target plated and dried naturally followed by addition of 1 µL α-Cyano-4-hydroxycinnamic acid (CHCA) matrix. Data were processed using AutoFlex software (Bruker Daltonik, Bremen, Germany).

7.2.5 Sample preparation for Circular Dichroism (CD) study

7.2.5.1 General

Consumables and reagents and preparation of solutions required for CD analysis are listed in **Table 7.1**.

7.2.5.2 Preparation of working solutions of hairpin oligonucleotides

5 µL of hairpin forming oligonucleotide was annealed in Eppendorf tubes as described in Section 7.2.1.2 in detail. To obtain DNA working solutions of 5 µM in a final volume of 1000 µL 995 µL of 100 mM ammonium acetate were added and the resulting solution homogenised for 5 s using a vortex mixer.

7.2.5.3 Preparation of ligand/DNA complexes

2 µL of 10 mM SJG-136 stock solution was added to 1000 µL 5 µM oligonucleotide solution resulting in a 4:1 (drug/DNA) ratio and incubated for 24 hours at 25 °C in a heating/cooling block. For t = 0 h experiments 2 µL of 10 mM SJG-136 stock solution was added to 1000 µL of 5 µM DNA working solution and subjected immediately to CD analysis as described below.

7.2.6 CD analysis

The CD spectra of the hairpin forming oligonucleotides and oligonucleotides/ligand complexes were acquired on a Chirascan plus spectrometer (Applied Photophysics Limited, Leatherhead, UK). The instrument was flushed continuously with pure evaporated nitrogen throughout the experiment. The CD spectra were measured between 200-400 nm in a strain-free rectangular 10 mm cell. Spectra were recorded using a 1 nm step size, a 1 s time-per-point and a spectral bandwidth of 1 nm. All spectra were recorded at 23 °C and the baseline was buffer (100 mM/20 mM ammonium acetate) corrected. Oligonucleotides were initially measured alone at a final concentration of 5 μM in an appropriate buffer. SJG-136 was added from a 10 mM stock solution in DMSO for a final concentration of 20 μM for a t = 0 experiment and CD spectra were recorded immediately followed by analysis of 24 hours incubation samples. Measurement was carried out while maintaining a constant concentration of oligonucleotides. Data were processed using OriginPro software (version 7.0). Differences in CD signals were plotted against the wavelength.

7.2.7 Sample preparation for Fluorescence Resonance Energy Transfer (FRET) DNA Melting assay

7.2.7.1 General

Hairpin forming oligonucleotides (**Figure 3.8**, Chapter 3) were fluorescence tagged with FAM (6-carboxyfluorescein) on 5'-end and TAMRA (6-carboxytetramethylrhodamine) on 3'- end. The sequences were purchased in lyophilised form from Eurogentec, Southampton, UK. Consumables and reagents including their manufacturer used for FRET experiments are listed in **Table 7.1**. Preparations of solutions required for FRET analysis are described in **Table 7.4** in detail. All prepared solutions were stored at room temperature. FRET buffer was stored at -20 °C and thawed to room temperature when required.

Table 7.4: Composition and preparation of solutions required for FRET analysis.

Solution	Composition
50 mM cacodylic acid	0.345 g cacodylic acid (137.99 g/mol) was dissolved in 20 mL distilled water. The volume was made up to 50 mL with distilled water.
FRET buffer	0.28 g KOH (56.11 g/mol) (100 mM) and 0.37 g KCl (74.55 g/mol) (100 mM) were dissolved in 20 mL distilled water. The volume was made up to 50 mL. pH was adjusted to 7.4 using 50 mM cacodylic acid.

7.2.7.2 Preparation of working solutions of hairpin oligonucleotides

Lyophilised hairpin forming oligonucleotides were dissolved in DEPC (diethylpolycarbonate)-treated water to form a stock solution of 100 µM. 20 µL of DNA stock solution was mixed with 80 µL FRET buffer to obtain a DNA solution of 20 µM. 20 µL of 20 µM DNA solution was mixed with 980 µL FRET buffer to form a working solution of 400 nM of each hairpin forming oligonucleotide. Fresh prepared working solutions were annealed by heating to 85 °C for 5 min in Eppendorf tubes in a heating/cooling block before cooling down to room temperature over 5 hours. All DNA solutions were stored at -20 °C and thawed to room temperature when required.

7.2.7.3 Preparation of working solutions of the PBD dimer SJG-136

Ligand working solutions were prepared by diluting a 10 mM SJG-136 DMSO stock solution in FRET buffer to double the required final concentration of 100 µM, 50 µM, 10 µM, 1 µM, 0.1 µM, and 0.01 µM to a final volume of 500 µL. Working solutions of SJG-136 were prepared freshly prior to FRET analysis.

7.2.7.4 Preparation of ligand/DNA complexes

25 µl of each annealed DNA working solution of 400 nM was pipetted in triplicates into a 96-well PCR multi plate (Bio-Rad, Hemel Hempstead, UK) followed by addition of 25µl of SJG-136 working solution in the desired concentration. This resulted in DNA final concentration of 200 nM. PCR plates were covered with Optical Flat 8-cap strips (Bio-Rad, Hemel Hempstead, UK) and incubated directly in the 96-well plate for 24 hours in the dark. Prior to FRET experiments 25 µL of each DNA working solutions of 400 nM was mixed with 25 µL FRET buffer and added in triplicates into the 96-well plate for background measurement.

	1	2	3	4	5	6	7	8	9	10	11	12
A	a	a	a	a	a	a	a	a	a			
B	b	b	b	b	b	b	b	b	b			
C	c	c	c	c	c	c	c	c	c			
D	d	d	d	d	d	d	d	d	d			
E	a	a	a	a	a	a						
F	b	b	b	b	b	b						
G	c	c	c	c	c	c						
H	d	d	d	d	d	d						

Figure 7.2: An example of a typical plate layout used for FRET analysis. a = representing annealed DNA mixed with FRET buffer for background measurement. b – d representing different concentrations of SJG-136 tested in the FRET assay.

NF-κB-1; NF-κB-2; EGR-1; AP-1; STAT3

7.2.8 FRET assay

FRET analysis was performed on a DNA Engine Opticon2 (Bio-Rad, Hemel Hempstead, UK) using the following cycling conditions: incubation at 30 °C for 5 s, fluorescence measurement every 0.5 °C for 30 s while heating from 30-100 °C, and finally incubation at 100 °C for 2 min. Fluorescence was captured at 515-545 nm. Obtained data were processed using OriginPro software (version 7.0, Origin Lab Corp.). The differences in

the melting temperature (ΔT_m) between adduct solution and blank (DNA sequences alone) were plotted against SJG-136 concentration using GraphPad Prism software Version 5.0 (GraphPad, California, USA).

7.2.9 Molecular modeling methodology

Molecular dynamics simulations were undertaken on SJG-136 non-covalently and covalently bound to the transcription factor consensus sequences used in the study. The DNA sequence in every case (including the TTT-loop) was constructed using the AMBER[299] module *nab*. The TTT-loop was then covalently linked to the backbone of DNA using parameters derived in-house. SJG-136 was docked in the minor groove using AMBER *xleap*, parm99SB and modified parmbsc0[300] and Gaff AMBER force field parameters. *Antechamber* was used to construct .mol2 files through the addition of Gasteiger charges and missing parameters were generated using *parmchk*. A covalent bond was generated between exocyclic amine groups of selected guanines (guided by molecular mechanics calculations[301]), forming either mono-adducts or inter/intra-strand cross-links. Energy minimisation was then undertaken in a gradient manner by initially placing the DNA under a high force constraint (*i.e.* 500 kcal mol^{-1} $Å^{-2}$), and this was then reduced in stages to zero to enable the ligand to find its local energy minimum, followed by reduction in force in a periodic manner with a relaxation of restraints. Production simulations in an implicit solvent (GBSA) were run for a period of 10 ns, and atomic coordinates were saved at 1ps intervals. Analysis of molecular dynamics simulations was undertaken using *VMD*[302] and *ptraj*, and all models were created using *Chimera*[303].

7.3 Experimental (biological evaluation of SJG-136)

7.3.1 Cell culture

7.3.1.1 General

The human breast cancer cell line MDA-MB-231 and the human colon cancer cell line HT-29 were maintained as described in Section 7.3.1.2. Consumables and reagents used during cell culture are listed in **Table 7.1**. All solutions were prepared in a class II microbiological safety cabinet (MSC). Commonly used solutions during cell culture are outlined in **Table 7.5**.

Table 7.5: Media preparation used during cell culture MDA-MB-231 and HT-29 human tumour cell lines.

Name	**Ingredients**
10% FBS media for MDA-MB-231	DMEM-Glautamax media was supplemented with 10% FBS, 1% MEM NEAA and 1% Penstrep solution. The media was mixed and stored at 2-8 °C.
10% FBS media for HT-29	Mc Coy's 5A media was supplemented with 10% FBS, 1% MEM NEAA and 1% Penstrep solution. The media was mixed and stored at 2-8 °C.

7.3.1.2 Maintenance and culture of cell lines

Unless otherwise stated, cells were maintained in 15 mL cell-line appropriate media containing 10% FBS in T75 flasks (75 cm^2) at 37 °C and 5% CO_2 in a Sanyo CO_2 incubator. Cell passaging was carried out within a Class II MSC. All solutions were pre-warmed to 37 °C in a water bath prior to use. Media was replaced every second day by aspiration, followed by addition of 15 mL fresh media.

7.3.1.3 Cell trypsinisation, centrifugation and counting

When cells reached approximately 70% confluence, media was aspirated and cells were washed by addition of 4 mL PBS. The PBS was aspirated and 2 mL of 0.25% trypsin-EDTA was added into the flask followed by incubation at 37 °C for 2 min. Cells were removed from the growing surface by tapping the flask on each side with hands. Trypsin-EDTA was inactivated and neutralised using 8 mL of appropriate growth media and pipetting up and down several times over the flask bottom to detach remaining adherent cells. The resulting suspension was transferred into a 20 mL universal falcon and centrifuged for 5 minutes at 21 °C and 1,500 rpm (Allegra™ X-22R centrifuge, Beckman Coulter, California, USA). The supernatant was aspirated, discarded and the resulting cell pellet re-suspended by addition of 10 mL appropriate media straight into the falcon followed by pipetting up and down to ensure re-suspension of the cell pellet. Viable cell number was counted using trypan blue exclusion method. A 1:1 trypan blue: cell suspension solution was prepared by mixing 100 µL of the cell suspension with 100 µL trypan blue, mixed and pipetted onto a haemocytometer. Due to the damaged cell membrane in non-viable cells, they are not able to remove the dye *via* exocytose and are therefore stained blue. Viable cells however, can exclude the trypan blue because of their intact cell membrane and are consequently not stained. The number of cells per mL of solution was calculated by counting all viable cells within 8 randomly selected grids of the haemocytometer under a microscope at 10 x magnification. The number of viable cells per ml was calculated using the following equation:

Cells/mL = (total cells counted/8) x 20,000

The cell suspension was diluted to the desired cell densities with appropriate media (stated in the individual assay descriptions) and plated in 6-well plates for polymerase chain reaction or T75 flasks western blot analysis.

7.3.1.4 Sample preparation for polymerase chain reaction (PCR) and Western blotting

Prior to RNA or protein extraction, cells were trypsinised and centrifuged as described in Section 7.3.1.3. The supernatant was aspirated, discarded and the resulting cell pellet was re-suspended in 10 mL of PBS followed by centrifugation for 5 min at 21 °C and 1,500

rpm in order to remove remaining media. The supernatant was aspirated, discarded and the cell pellets snap-frozen in liquid nitrogen. Frozen pellets were stored at -80 °C until PCR or Western blot analysis.

7.3.2 Reverse Transcription (RT-PCR)

7.3.2.1 General

All isolation, RT-PCR and PCR sample preparation steps were performed in a PCR hood (SLEE, Mainz, Germany). Consumables and reagents used during RT-PCR are listed in **Table 7.1**.

7.3.2.2 RNA isolation from frozen cell pellets

Total RNA was extracted from snap-frozen cell pellets using the RNeasy® Plus Mini Kit according to the manufacturer's instructions. Briefly, the frozen cell pellet was re-suspended in 350 µL RLT Plus lysis buffer, pipetted up and down multiple times and homogenised for 30 s using a vortex mixer. The lysate was transferred into a gDNA Eliminator spin column placed in 2 mL collection tube and centrifuged for 30 s at 8,000 g in a Sigma 1-14 Microcentrifuge (Appleton Woods, Birmingham, UK) in order to remove remaining genomic DNA. The gDNA eliminator spin column was discarded, the flow-through mixed with 350 µL 70% ethanol, and mixed by pipetting up and down 10 times. The entire volume of 700 µL was transferred into RNeasy spin column placed in a 2 mL collection tube and spun for 15 s at 8,000 g in order to bind the RNA to the silica membrane. The flow-through was discarded and 700 µL of buffer RW1 were added to the RNeasy spin column placed in a new 2 mL collection tube followed by centrifugation at 8,000 g for 15 s. Next, the flow-through was again discarded and 500 µL of buffer RPE were added into the RNeasy spin column. After centrifugation at 8,000 g for 15 s, the flow-through was removed and the wash step with 500 µL buffer RPE repeated with centrifugation at 8,000 g for 2 min. Finally, the RNeasy spin column was placed in a new 1.5 mL collection tube and 50 µL of RNase-free water was added directly to the spin column membrane, followed by centrifugation at 8,000 g for 1 min. The isolated RNA was stored for maximum two weeks at -80 °C until reverse transcription analysis.

7.3.2.3 RT of mRNA

Isolated mRNA was reverse-transcribed into cDNA using the High Capacity RNA-to-cDNA as per manufacturer's instructions. Briefly, a mixture was prepared consisting of 10 µL 2 X RT buffer containing dNTPs, random octamers, oligodT-16, 1 µL 20 X enzyme mix containing MuLV (Murine Leukaemia Virus) reverse transcriptase and RNase inhibitor protein and 9 µL extracted RNA. A negative RT control was prepared following the same procedure. 1 µL 20 X enzyme mix was replaced with 1 µL nuclease-free water to ensure samples do not contain remaining genomic DNA. Each sample tube was homogenised for 5 s and spun briefly prior to transfer into a PCR Sprint thermal cycler (Thermo Scientific, Loughborough, UK). Incubation was carried out at 37 °C for 1 hour followed by 5 min at 94 °C in order to terminate the reaction. Complementary DNA samples were stored at -20 °C prior to PCR analysis.

7.3.3 PCR

7.3.3.1 General

Primers used for PCR experiments were purchased in lyophilised form from Integrated DNA Technologies, Coralville, USA. Primer preparation for PCR is described below (Section 7.3.3.2) in detail. All PCR primer and sample preparation steps were performed in a PCR hood (SLEE, Mainz, Germany). Consumables and reagents used during PCR analysis including their manufacturer are listed in **Table 7.1**.

7.3.3.2 Primers preparation for PCR experiments

Specific primers used for genes of interest are listed in **Table 7.6** and **7.7**. Prior to use, lyophilised primers were centrifuged for 5 s and diluted with PCR grade water to form stock solutions of 100 µM. For PCR experiments appropriate primer pairs were further diluted 1:10 with PCR grade water to form working solutions of 10 µM. Stock and working solutions were stored at -20 °C, thawed and vortexed prior to use.

Table 7.6: Primers used during PCR for expression of STAT3 dependent genes. F = Forward; R = Reverse.

Gene	Sequence
Bcl-2	F: 5'-CAT CTT CTC CTC CCA GCC C-3' R: 5'-CCG AAC TCA AAG AAG GCC AC-3'
Cyclin D1	F: 5'-ACA GAT CAT CCG CAA ACA CG-3' R: 5'-CTC CTC CTC TTC CTC CTC CT-3'
NNMT	F: 5'-TGGCCCCACTATCTATCAGC-3' R: 5'-CCTCTTTCACAGCAGCCTCT-3'
STAT3	F: 5'-CTT TGA GAC CGA GGT GTA TCA CC-3' R: 5'-GGT CAG CAT GTT GTA CCA CAG G-3'
STAT1	F: 5'-CTA GTG GAG TGG AAG CGG AG-3' R: 5'-CAC CAC AAA CGA GCT CTG AA-3'
Fascin	F: 5'-ACT GGC TAC ACG CTG GAG TT-3' R: 5'-GGA AGG CAC ACT TTT TGG TG-3'
GAPDH	F: 5'-AGC CAC ATC GCT CAG ACA C-3' R: 5'-GCC CAA TAC GAC CAA ATC C-3'

Table 7.7: Primers used during RT-PCR for expression of AP-1 dependent genes. F = Forward; R = Reverse.

Gene	Sequence
VEGF	F: 5'-CTA CCT CCA CCA TGC CAA GT-3' R: 5'-GGT GCT CCA GAA GTG AAT GC-3'
Cyclin D1	F: 5'-CCC TCG GTG TCC TAC TTC AA-3' R: 5'-GGC GGA TTG GAA ATG AAC T-3'
CREB5	F: 5'-ATT CCC ACC TTC ATG CAC A-3' R: 5'-TTG TGT TGC TGG TGA AAC CT-3'
Elk-1	F: 5'-GCT TCC TAC GCA TAC ATT GAC C -3' R: 5'-GGT GCT CCA GAA GTG AAT GC-3'
AP-1	F: 5'-CCA AAG GAT AGT GCG ATG TTT-3' R: 5'-CTG TCC CTC TCC ACT GCA AC-3'
p53	F: 5'-AGG CCT TGG AAC TCA AGG AT-3' R: 5'-CCC TTT TTG GAC TTC AGG TG-3'
GAPDH	F: 5'-AGC CAC ATC GCT CAG ACA C-3' R: 5'-GCC CAA TAC GAC CAA ATC C-3'

7.3.3.3 PCR amplification of cDNA

Obtained cDNA samples from RT-PCR reaction were prepared as followed: 24 µL of a master mixture consisting of 2.5 µL 10 X PCR reaction buffer, 0.5 µL 10 mM deoxyribonucleotide triphosphate bases (dNTP mix, 0.2 mM final concentration), 0.75 µL 50 mM $MgCl_2$ (1.5 mM final concentration), 0.625 µL of 10 µM forward primer (250 nM final concentration), 0.625 µL of 10 µM reverse primer (250 nM final concentration), 0.1 µL of 5 U/µL platinum Taq polymerase (0.02 U/µL final concentration), and 18.9 µL water PCR grade. 24 µL of the master mix was added to 1 µL cDNA sample and mixed up and down several times. PCR was performed in a PCR Sprint thermal cycler (Thermo Scientific, Loughborough, UK) using the following cycle conditions: 94 °C for 5 min, 30 cycles of 94 °C for 30 s, 55 °C for 30 s, and finally 72 °C for 1 min, with a final extension at 72 °C for 5 min. For each gene a *GAPDH* control was prepared and run under the same

conditions. Once the reaction was complete, the samples were held at 4 °C.

7.3.4 Agarose gel electrophoresis

7.3.4.1 General

Consumables and reagents used during agarose gel electrophoresis including their manufacturer are listed in **Table 7.1**.

7.3.4.2 Agarose gel electrophoresis of amplified DNA

Confirmation of target gene amplification was carried out using agarose gel electrophoresis. 0.75 g agarose was dissolved in 50 mL 1 X TAE buffer (prepared by mixing 30 mL of 10 X TAE stock solution with 270 mL deionised water) in a conical flask and heated in a microwave for 1 min and 10 s. The heating process was interrupted every 20 s and the flask mixed gently to ensure homogenisation of the agarose in the buffer. While agarose solution cooled down, a 12-well gel comb was placed into the gel tray. 2.5 µL of 10,000 x GelRed were added into the agarose solution and mixed before poured into the tray and left to set for approximately 30-40 min. 5 µL of each PCR sample were mixed with 1 µL 6 X blue gel loading buffer and loaded into a well of the agarose gel submerged in TAE buffer. 1 µL of 100 bp ladder were mixed with 1 µL loading buffer and 4 µL deionised water before 2 µL were loaded to the leftmost well of the gel. Electrophoresis was performed at 50 V for 40 min and visualised at 300 nm using UV transillumination (U:GENIUS, Syngene, Cambridge, UK) and then digitally scanned. Band size was determined by comparison with a 100 bp ladder. Amplified PCR samples were stored at -20 °C.

7.3.5 Quantitative PCR (qPCR)

7.3.5.1 General

Primers used for qPCR were purchased in lyophilised form from Integrated DNA Technologies, Coralville, USA. UPL probes were supplied by the Genomic Centre, Waterloo Campus, King's College London.

7.3.5.2 Primers preparation for qPCR experiments

Preparation of primer pairs specific genes of interest for qPCR analysis is described in Section 7.3.3.2 in detail. Primers used in the qPCR study are listed in **Table 7.8**.

Table 7.8: Primers used during qPCR for expression of STAT3 and AP-1 dependent genes. F = Forward; R = Reverse.

Gene Name	Accession Number	Primer Sequence	Open Reading Frame	Amplicon Size (nt)	UPL
Bcl-2	ENSG 00000171791	F: 5'-AGT ACC TGA ACC GGC ACC T-3' R: 5'-GCC GTA CAG TTC CAC AAA GG-3'	1998– 2016 2052 - 2071	74	75
Cyclin D1	ENSG 00000110092	F: 5'-CCC TCG GTG TCC TAC TTC AA-3' R: 5'-GGC GGA TTG GAA ATG AAC T-3'	226 – 245 313 - 331	106	67
Survivin	NM_001168	F: 5'-TCT GCT TCA AGG AGC TGG A-3' R: 5'-AAA GTG CTG GTA TTA CAG GCG TA-3'	202 – 220 267 - 289	88	36
STAT3	ENSG 00000168610	F: 5'-GAG CAG AGA TGT GGG AAT GG-3' R: 5'-CGG TCT CAA AGG TGA TCA GG-3'	1375– 1394 1443 - 1462	88	17
VEGF	ENSG 00000112715.16	F: 5'-CTA CCT CCA CCA TGC CAA GT-3' R: 5'-GGT GCT CCA GAA GTG AAT GC-3'	1111– 1130 1176 - 1196	86	29
CREB5	ENSG 00000146592.11	F: 5'-ATT CCC ACC TTC ATG CAC A-3' R: 5'-TTG TGT TGC TGG TGA AAC CT-3'	996 – 1014 1071 - 1090	95	9
Elk-1	ENSG 00000126767.11	F: 5'-GCT TCC TAC GCA TAC ATT GAC C-3' R: 5'-GGT GCT CCA GAA GTG AAT GC-3'	1168– 1189 1224 - 1243	76	16
AP-1	ENSG 00000177606.5	F: 5'-CCA AAG GAT AGT GCG ATG TTT-3' R: 5'-CTG TCC CTC TCC ACT GCA AC-3'	3012– 3032 3054 - 3073	62	19
p53	ENSG 00000141510.8	F: 5'-AGG CCT TGG AAC TCA AGG AT-3' R: 5'-CCC TTT TTG GAC TTC AGG TG-3'	979 – 998 1044 - 1063	85	12
ß-actin	NM_001101.3	F: 5'-CCA ACC GCG AGA AGA TGA-3' R: 5'-CCA GAG GCG TAC AGG GAT AG-3'	425 - 442 502 - 521	97	64

7.3.5.3 qPCR of cDNA

qPCR was used to quantify changes in gene expression in the human cancer cell lines MDA-MB-231 and HT-29. Data for qPCR analysis were collected from three independently performed assay rounds for each cell line. Primers were tested with non-quantitative PCR to optimise reaction conditions to generate a single PCR product, as confirmed by gel electrophoresis.

Briefly, 18 µL of each resulting cDNA sample including no RT control were pipetted onto a well of a 96-well PCR plate and diluted to 200 µL by the addition of 182 µL DEPC water. A pooled reference control DNA was used as a quality control for plate – to – plate variation. The pool was created by mixing 2 µL of each cDNA sample and dilution to 200 µL with DEPC (diethylpolycarbonate)-treated water. From this pool three different dilutions 1:10, 1:100, and 1:1000 were prepared to create a dilution series that covered the entire range of concentration measured in the assay. 1:10 dilution was prepared by removing 20 µL of pool solution and pipetting into a new well followed by addition of 180 µL of DEPC water. 1:100 dilution was created by removing 20 µL of 1:10 dilution and pipetting into a new well followed by addition of 180 µL of DEPC water. 1:1,000 dilution was prepared by removing 20 µL of 1:100 dilution and pipetting into a new well followed by addition of 180 µL of DEPC water. The last well of the dilution series contained 200 µL DEPC water for background measurement. The assay mix was prepared in Eppendorf tubes for each gene separately by mixing 953.01 µL of fast start universal probe master mix, 76.245 µL of 10 µM forward primer, 76.245 µL of 10 µM reverse primer, and 19.06 µL appropriate UPL probe. Primers and UPL probes used for qPCR experiments are listed in **Table 7.8**. The master mix was homogenised by pipetting up and down multiple times. 179 µL of appropriate master mix were pipetted into a deep well block according to sample number. 4 µL of diluted cDNA were combined with 6 µL of the assay mix in a 384-well plate on a Biomek FX liquid handling robot (Beckman Coulter, UK). Once the mixing process was finished, the plate was removed from the robot, sealed with a plastic lid and spun for 10 s to collect the liquid at the bottom of the 384-well plate. qPCR was performed on an ABI Prism 7000 thermal cycler (Applied Biosystems) using the following cycling conditions: 95 °C for 10 min, followed by 40 cycles at 95 °C for 15 s and 60 °C for 1 min. All samples were run in triplicate together with their corresponding negative control. Negative template controls and the pooled

reference control DNA were run on every 384-well plate. *ß-actin* was utilised as a reference gene to normalise mRNA expression. Results were analysed using the Data Assist Software Version 3.0 (Applied Biosystems) and gene expression levels were quantified using the comparative ΔΔCt method[6]. Results were expressed as ΔCt (Ct sample – Ct *ß–actin*) and the fold change in gene expression compared to the untreated sample.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A	1	1	1	1	1	1	Pool					
B	1a	1a	1a	1a	1a	1a	1:10					
C	2	2	2	2	2	2	1:100					
D	2a	2a	2a	2a	2a	2a	1:1,000					
E	3	3	3	3	3	3	water					
F	3a	3a	3a	3a	3a	3a						
G												
H												

Figure 7.3: A typical plate layout used for qPCR. 1 = untreated cells, 1a = no RT control of 1; 2 = stimulated cells, 2a = no RT control of 2; 3 = stimulated cells followed by drug treatment, 3a = no RT control of 3.

MDA-MB-231; **HT-29**

7.3.6 Western Blotting

7.3.6.1 General

Consumables and reagents including their manufacturer required for Western blotting are listed in **Table 7.1**. Preparation of solutions used for Western blotting is described in **Table 7.9** in detail.

Table 7.9: Buffer solutions and preparation used during Western blotting.

Name	**Ingredients**
RIPA buffer (Radioimmunioprecipitation buffer)	5 mL of 1 M (157.6 g/L in water) Trizima HCl (50 mM, pH 7.4), 1 mL of 1% (v/v) Nonidet P40 (0.1%), 0.5 g 0.5% sodium deoxycholate (0.0025%), 15 mL of 1 M (58.4 g/L in water) NaCl (150 mM), 0.15 g EDTA (5 mM), 0.1 g 0.1% (w/v) SDS (0.0001%) were mixed and made up to 100 mL deionised water. The buffer was aliquoted followed by addition of complete Protease Inhibitor Cocktail and PhosSTOP tablet to each aliquot. RIPA buffer was stored at -20 °C and thawed prior to use.
TBS (Tris Buffered Saline)	6 g Trizima HCl (50 mM Tris) and 12 g NaCl (200 mM) were dissolved in 500 mL deionised water, adjusted to pH 7.4 and made up to 1 L with deionised water. TBS buffer was stored at room temperature.
TBST (Tris Buffered Saline – Triton X – 100)	500 µL of Triton X-100 was added to 1 L TBS for a final concentration of 0.05%. TBST buffer was stored at room temperature.
MTBST (Milk - Tris Buffered Saline – Triton X – 100)	50 g of skimmed milk powder was dissolved in 1 L TBST for a final concentration of 5%. MTBST solution was prepared fresh when required.
AB I (Anode buffer I)	18.15 g Tris base (0.3 M) was dissolved in 50 mL methanol before 400 mL of deionised water were added. pH was

	adjusted to 10.4 and solution was made up to 500 mL using deionised water. AB I was stored at room temperature.
AB II (Anode buffer II)	1.51 g Tris base (25 mM) was dissolved in 50 mL methanol before 400 mL of deionised water were added. pH was adjusted to 10.4 and solution was made up to 500 mL using deionised water. AB II was stored at room temperature.
CB (Cathode buffer)	1.51 g Tris base (25 mM) and 1.5 g glycine (40 mM) were dissolved in 50 mL methanol followed by addition of 400 mL of deionised water. pH was adjusted to 9.4 and solution was made up to 500 mL using deionised water. CB was stored at room temperature.

7.3.6.2 Sample preparation for protein assay

Snap-frozen cell pellets were lysed directly in 20 mL universal tubes using 100 µL RIPA buffer and vortexed for 30 s to improve lysis. The lysates were transferred into new Eppendorf tubes and centrifuged at 13,000 rpm for 10 min in a Sigma 1-14 Microcentrifuge in order to precipitate the cells. Supernatants were transferred into new Eppendorf tubes and two different dilutions (1:5 and 1:10) were prepared using RIPA buffer as described in Section 7.3.6.4 in detail. Samples not used for protein assay were kept on ice until further processing.

7.3.6.3 Protein assay

A protein assay was carried out prior to SDS PAGE in order to determine the total protein concentration in each sample to ensure that approximately same amount of total protein is loaded onto the gel. Briefly, lysed samples were prepared in labelled Eppendorf in a

1:5 dilution by mixing 6 µL original sample with 24 µL RIPA buffer resulting in a total volume of 30 µL. From this 1:5 dilution 10 µL were removed and mixed with 10 µL RIPA buffer in a new labelled Eppendorf tube to a final volume of 20 µL. Protein standards of 1.0 mg/mL, 0.8 mg/mL, 0.6 mg/mL, 0.4 mg/mL, 0.2 mg/mL, and 0.0 mg/mL were made from a 10 mg/mL stock solution of BSA prepared in RIPA buffer. Volumes of RIPA buffer and BSA stock solution required for the dilution series are described in **Table 7.10** in detail. The protein assay, based on the methodology of Lowry *et al.* 1951 was performed using Bio-Rad DC Protein Assay as per manufacturer's instructions. For one 96-well plate (Thermo Scientific, Loughborough, UK) 2 mL of Bio-Rad DC Protein Reagent A were mixed with 40 µL of Solution S resulting in Solution A. 5 µL of each BSA standard concentration and 5 µL of each sample dilution were pipetted in triplicates into a 96-well plate as outlined in **Figure 7.4**. 25 µL of Solution A were added with a multichannel pipette to each well and the plate was carefully vortexed covered with its lid. 200 µL of Bio-Rad Reagent B were added to each well using a multichannel pipette followed by homogenisation using the vortex mixer. The plate was incubated at room temperature in the dark for 15 min and existing bubbles were removed with a needle before the absorbance was read at 750 nm using a spectrophotometer (Wallac Victor 2, 1420 Multilabel Counter, PerkinElmer, Hopkinton, USA). The total protein concentration in each sample was determined based on the albumin calibration curve and represent in µg/µL.

Table 7.10: Volumes of BSA stock solution and RIPA lysis buffer used for protein standard preparation.

BSA concentration (mg/mL)	**BSA stock solution (µL)**	**RIPA lysis buffer (µL)**
1.0	10	90
0.8	8	92
0.6	6	94
0.4	4	96
0.2	2	98
0.0	0	100

	1	2	3	4	5	6	7	8	9	10	11	12
A	1.0	1.0	1.0	1:5	1:5	1:5	1:10	1:10	1:10	Sample 1		
B	0.8	0.8	0.8	1:5	1:5	1:5	1:10	1:10	1:10	Sample 2		
C	0.6	0.6	0.6	1:5	1:5	1:5	1:10	1:10	1:10	Sample 3		
D	0.4	0.4	0.4									
E	0.2	0.2	0.2									
F	0.0	0.0	0.0									
G												
H												

Figure 7.4: A typical 96-well plate layout used for protein assays. Numbers 1.0 – 0.0 represent BSA standard concentrations in mg/ mL in RIPA buffer. Sample 1 = untreated cells (MDA-MB-231/HT-29); sample 2 = stimulated cells (MDA-MB-231/HT-29); sample 3 = stimulated cells followed by drug treatment (MDA-MB-231/HT-29). 1:5 and 1:10 represent the dilution factor of the samples in RIPA buffer.

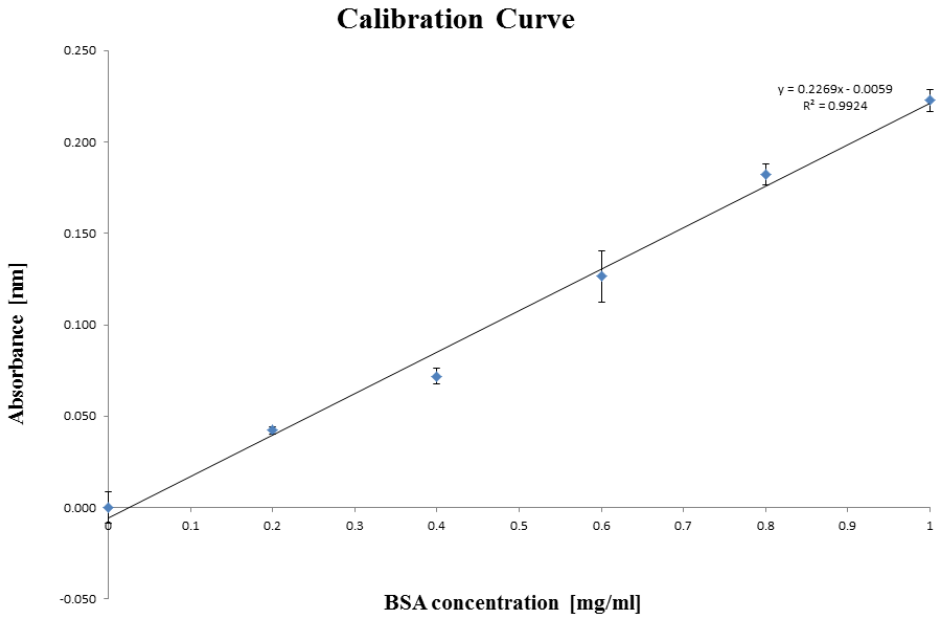


Figure 7.5: An example of a BSA standard calibration curve.

7.3.6.4 Sample preparation for Western blotting

Once the protein concentration was calculated a dilution using RIPA buffer was prepared to give a protein concentration of 3 µg/µL in each sample. Laemmli sample buffer was mixed with each sample in a 1:1 ratio for a final protein concentration of 1.5 µg/µL and heated for 5 min to 100 °C in an Eppendorf heater (Grant Instruments, Shepreth, UK). This boiling step aims to reduce disulfide linkages initiated by the reducing agent 2-mercaptoethanol contained in the Laemmli buffer. This results in breaking up the quaternary protein structure, and protein separation during gel electrophoresis occurs due to protein size.

7.3.6.5 SDS – polyacrylamide gel electrophoresis (PAGE)

SDS–PAGE was carried out using pre-cast 4–12% NuPage Bis–Tris SDS–PAGE gel submerged in 1 L 1 X MOPS (3-(*N*-morpholino)propanesulfonic acid) running buffer (prepared by mixing 50 mL 20 X MOPS running buffer with 950 mL deionised water). 60 µL of each sample (equivalent to 30 µg total protein) were loaded into the wells of the gel. For protein size comparison 5 µL of Novex Sharp Pre-stained Protein Standard were loaded to the leftmost well of the gel. Protein separation was performed at 200 V for 50 min.

7.3.6.6 Semi – dry transfer of proteins to nitrocellulose membrane

The semi-dry transfer onto a Hybond-ECL nitrocellulose membrane was carried out according to the semi-dry transfer protocol reported in Martin *et al.* 2013. Briefly, AB I, AB II and CB were poured into separate trays. Once SDS-PAGE was finished, the gel was carefully removed from the electrophoresis assembly followed by the opening of the gel plates. Wells and the left top corner were cut off in order to determine the gel orientation. Gel was soaked in CB and placed on a rocking platform (Labnet International, New Jersey, USA) for minimum 10 min to equilibrate the gel with the buffer. Hybond-ECL nitrocellulose membrane was cut to the size of the Quick Draw Filter (QDF) paper and soaked for 2 min in deionised water before being transferred into AB II and placed on the rocking platform for 5 min. Quick Draw Filter paper was soaked in appropriate buffer and a gel/membrane "sandwich" was built as shown in **Figure 7.6** starting from

the bottom of the Trans-Blot® SD Semi-Dry Transfer Cell tank (Bio-Rad, Hemel Hertfordshire, UK). Once all the layers have been set up, transfer was performed at 19 V for 36 min if only one gel and for 42 min if two gels were being transferred.

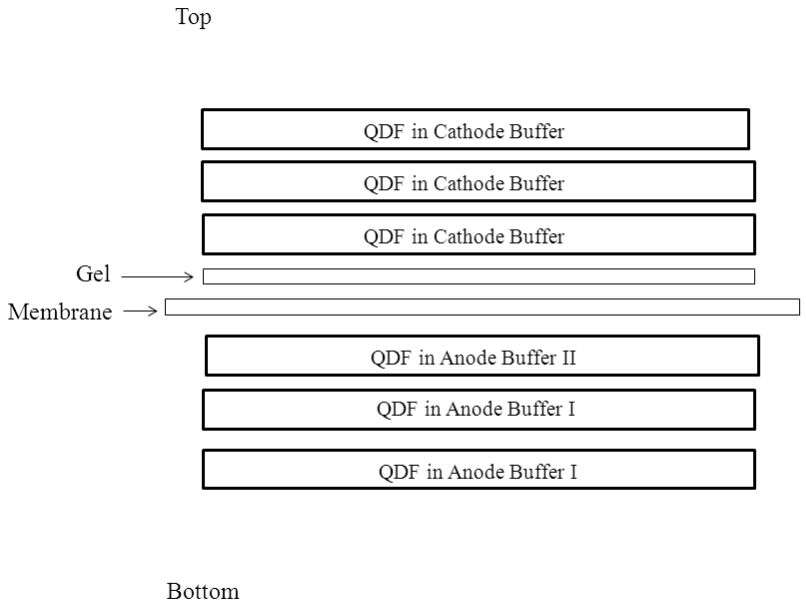


Figure 7.6: Gel/membrane "sandwich" composition used during the semi-dry transfer of proteins from a 4-12% NuPage Bis-Tris SDS-PAGE gel to a nitrocellulose membrane. QDF = Quick Draw Filter paper.

7.3.6.7 Immunoblotting

Once the transfer was finished, the gel/membrane sandwich was disassembled and the nitrocellulose membrane was quickly cut to the size of the gel. The membrane was transferred into a tray containing MTBST and blocked for at least 1 hour on a rocking platform at room temperature. This blocking step aims to circumvent non-specific binding which might occur due to interactions between the membrane and the antibody applied to detect the target protein resulting in "noise" and false positive results. The proteins contained in the blocking solution will bind to the membrane in all places where the proteins of interest have not attached. Consequently, the primary antibody will bind only to its target protein. Antibodies against the target proteins are listed in **Table 7.1** in Section 7.1. They were prepared in 50 mL falcon tubes by mixing with MTBST in 1:250 dilution and stored at -20 °C and thawed on ice prior to use. The membrane was

transferred straight into the falcon tube and incubated with primary antibodies overnight at 4 °C on a tube roller (Spiramix, Thermo Scientific, Loughborough, UK). On the next day, the blot was rinsed with MTBST at room temperature before being incubated for 2 hours in MTBST at room temperature on a rocking platform. The secondary antibody conjugated to horseradish peroxidase (HRP) was prepared in 20 mL universal tube by mixing with MTBS in a 1:1,000 dilution and addition of StrepTactin-HRP conjugate in a 1:4,000 ratio. The membrane was incubated with the secondary antibody mixture for 1 hour under the same conditions. The StrepTactin-HRP conjugate was added to the secondary antibody mix to enhance the signal of the Strep-tagged protein markers. The secondary antibody was discarded and the membrane was rinsed and washed with MTBST at room temperature on a rocking platform for 5 min. MTBST was removed and the blot washed twice in TBST for 5 min using the same conditions. Lastly, TBST was discarded and the membrane was submerged in TBS prior to detection. The ECL detection mixture was prepared by mixing equal volumes of solution A and B contained in the ECL detection kit as per manufacturer's instructions. The membrane was removed from TBS and slightly tapped with its corner on tissue paper in order to remove remaining TBS. ECL reagent was applied onto the membrane and incubated for 5 min at room temperature and dried as described above. Membranes were visualised using a ChemiDoc MP transilluminator (Bio-Rad, Hemel Hempstead, UK). To re-probe α-tubulin as loading control the blots were stripped with Restore™ western blot stripping buffer for 1 hour at room temperature on a rocking platform followed by rinsing and re-blocking for another hour. The immunoblotting process was repeated using α-tubulin as described above. Images were digitally captured and band intensities were quantified by densitometry utilising the ImageLab software (Version 4.1, Bio-Rad, Hertfordshire, UK) and normalised for α-tubulin. Protein quantification was carried out by dividing the integrated peak area of the protein of interest by the integrated peak area of the reference gene α-tubulin. Alternatively, blots were captured using the traditional X-ray film development. This developing procedure was carried out in a dark room. ECL detection kit was prepared and applied onto the nitrocellulose membrane as described above. The blot was placed in a developing cassette and closed with a thin film covered by an X-ray film. The cassette was closed for approximately 1 min in order to press the membrane onto the film. The cassette was opened carefully, the film directly placed into the film developer solution and developed until protein bands appeared. Once bands were visible, the film was transferred into the fixation solution for approximately 5 min and finally placed in

water for wash. The developed films were air-dried overnight and scanned before quantified by densitometry using ImageJ software (version 1.49, National Institutes of Health, USA) and normalised for α-tubulin expression. Quantification of band intensity was performed as described above.

7.3.7 Statistical analysis

Statistical analysis was performed using GraphPad Prism software (GraphPad, California, USA). *P* values < 0.05 were considered as significant.

Appendix Chapter 2: Additional HPLC chromatograms and MALDI-TOF spectra

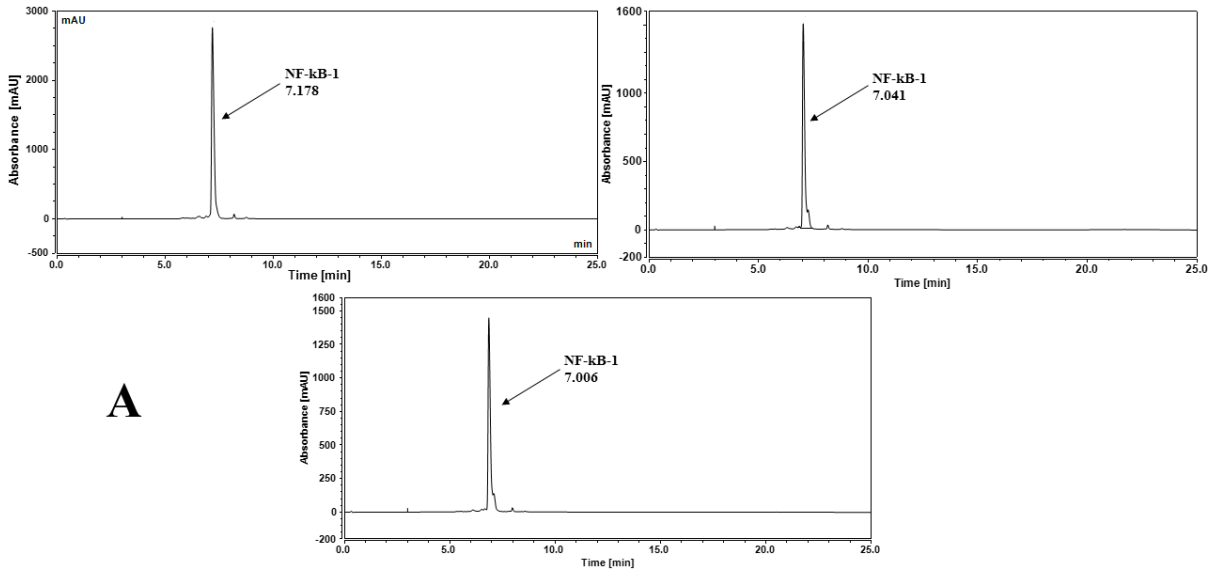


Figure A2.1A: HPLC chromatograms and the corresponding RTs for the sequence NF-κB-1 after three injections. Refers to Chapter 2, p 96 in thesis.

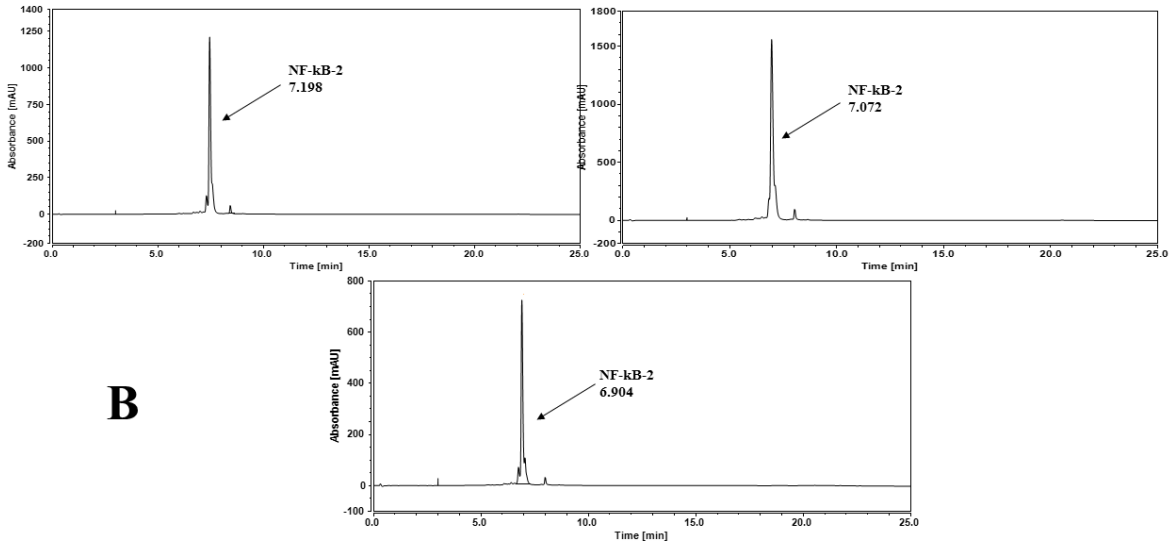
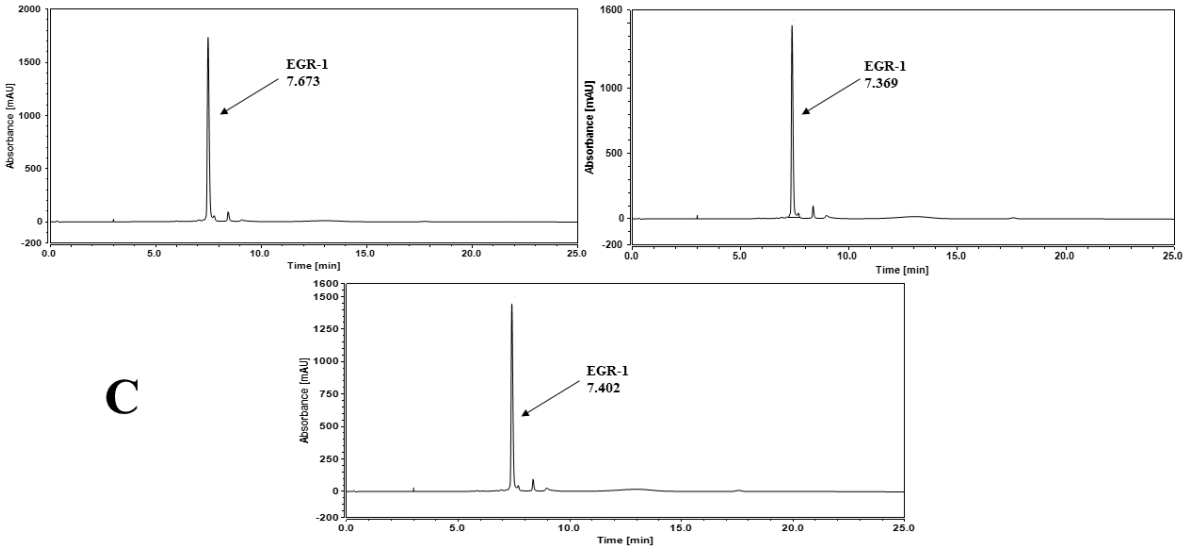
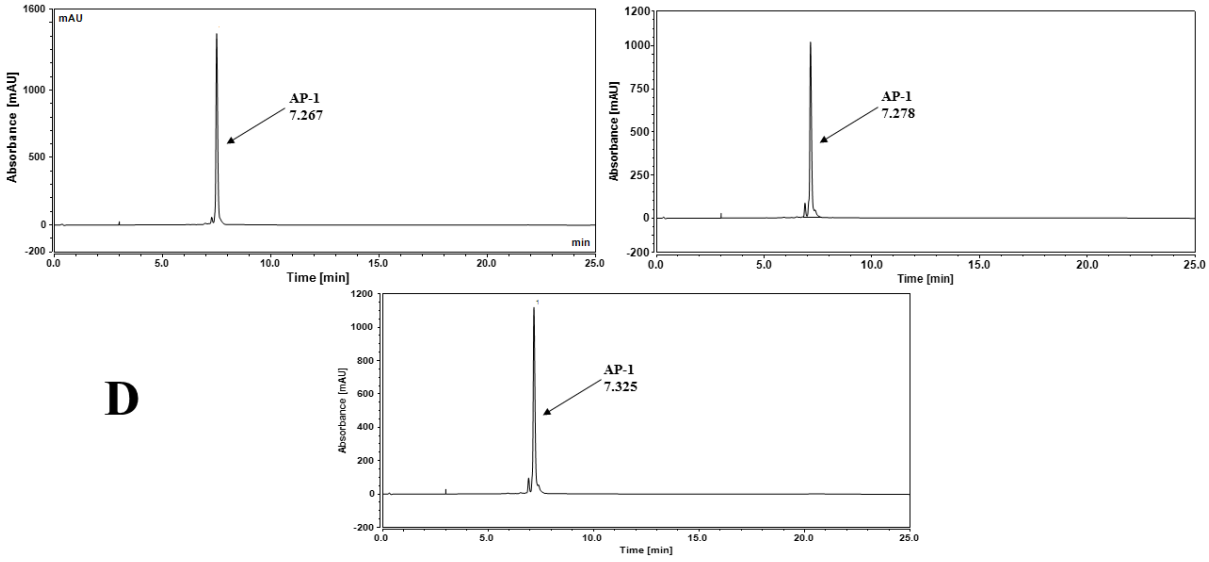


Figure A2.1B: HPLC chromatograms and the corresponding RTs for the sequence NF-κB-2 after three injections. Refers to Chapter 2, p 96 in thesis.



C

Figure A2.1C: HPLC chromatograms and the corresponding RTs for the sequence EGR-1 after three injections. Refers to Chapter 2, p 96 in thesis.



D

Figure A2.1D: HPLC chromatograms and the corresponding RTs for the sequence AP-1 after three injections. Refers to Chapter 2, p 96 in thesis.

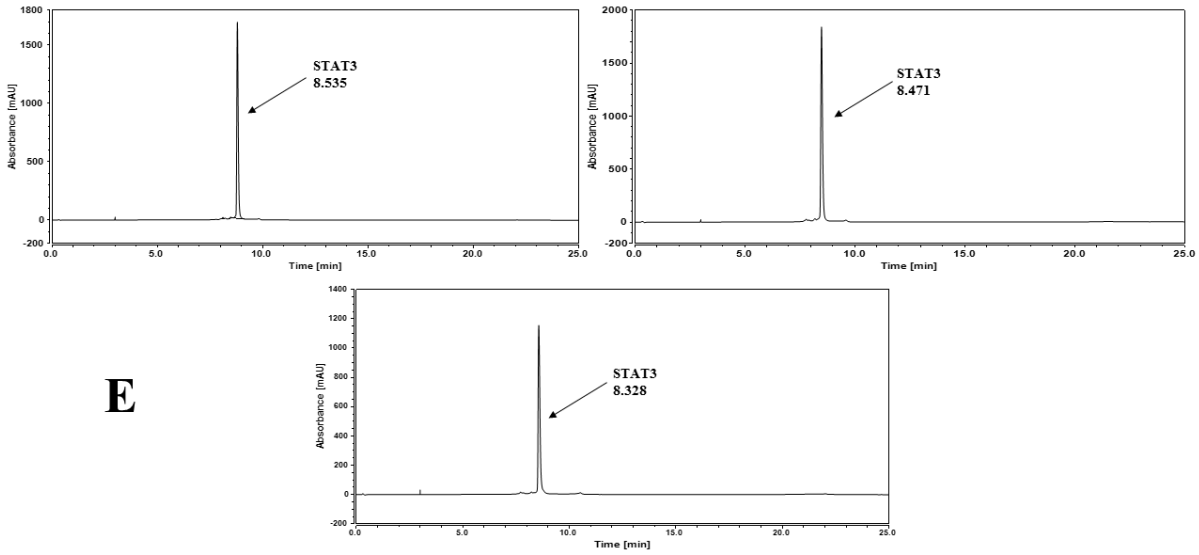


Figure A2.1E: HPLC chromatograms and the corresponding RTs for the sequence STAT3 after three injections. Refers to Chapter 2, p 96 in thesis.

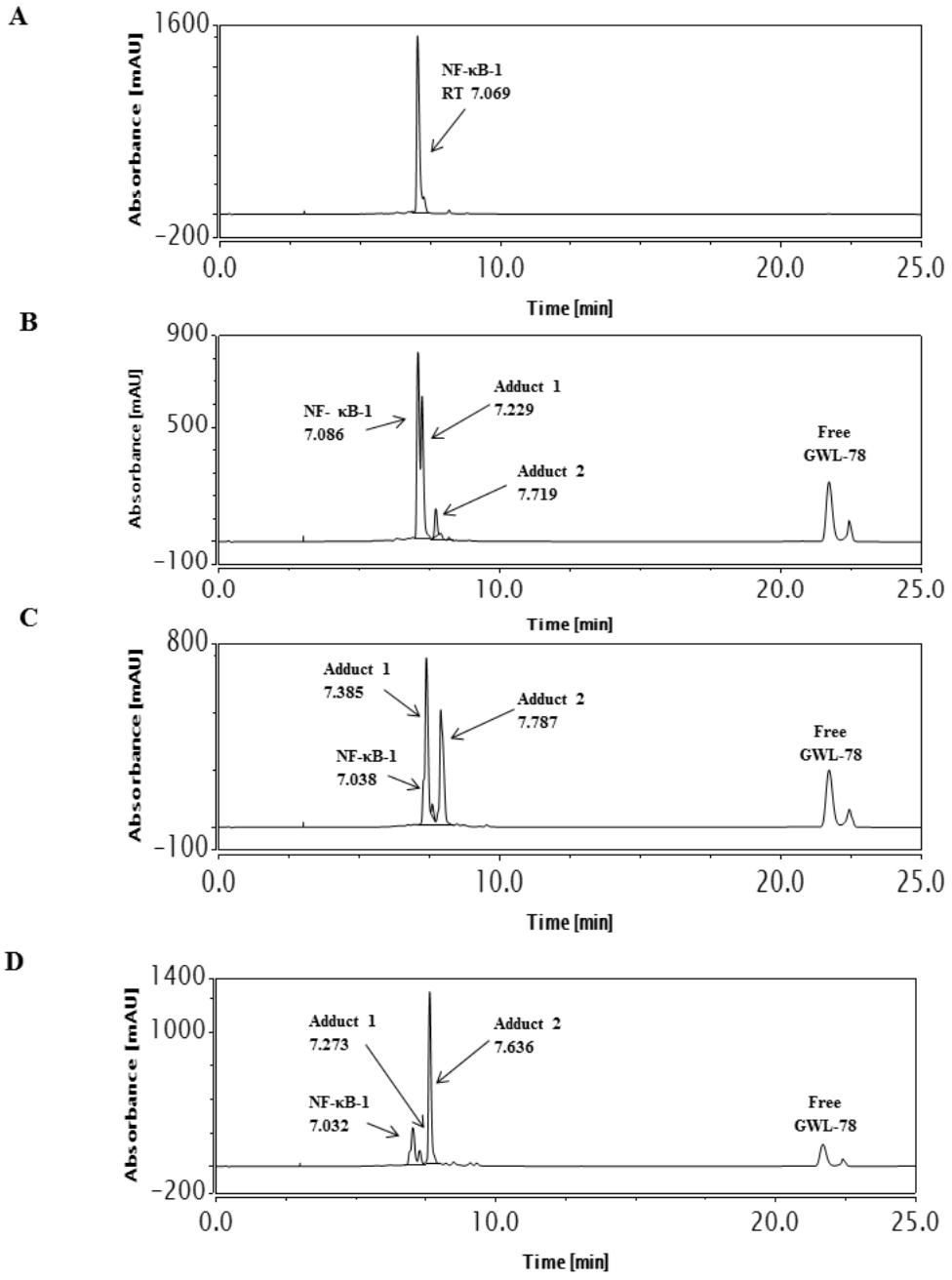


Figure A2.2: Time-course HPLC chromatograms: (**A**) Annealed NF-κB-1 sequence; (**B**) Annealed NF-κB-1 sequence after incubating with GWL-78 for 5 min showing the appearance of two new peaks at RT 7.229 min (**adduct 1**) and RT 7.719 min (**adduct 2**); (**C**) after 3 hours incubation; (**D**) after 24 hours incubation showing increase of **adduct 2** formation with reaction not complete after 24 hours. Refers to Chapter 2, p 97 in thesis.

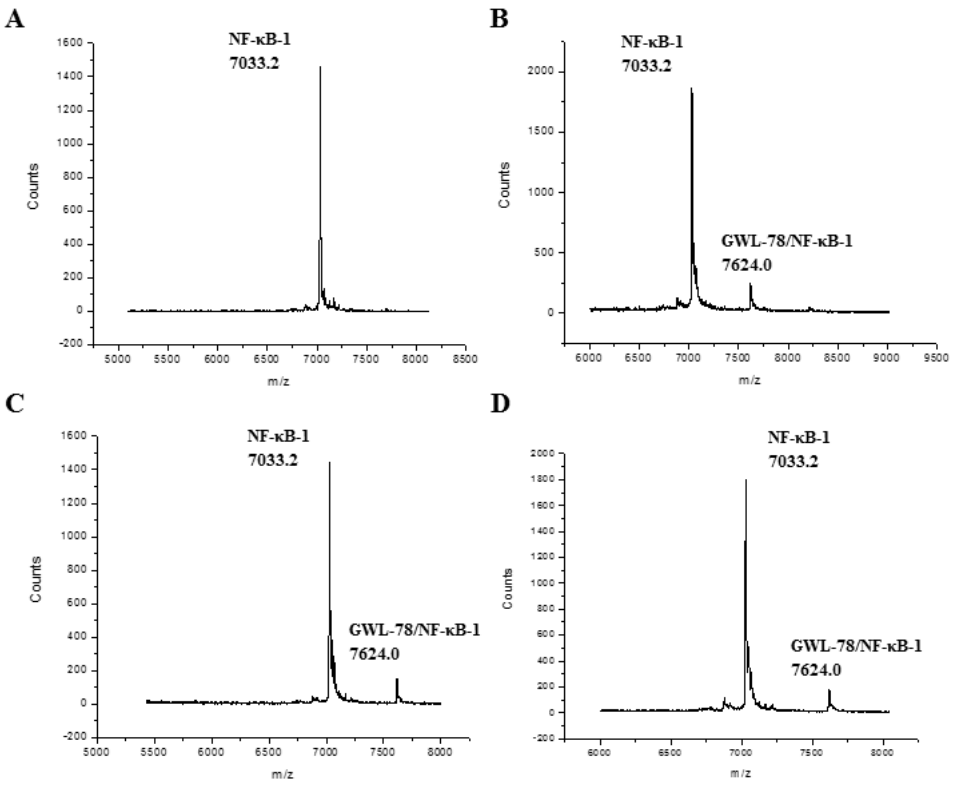


Figure A2.3: MALDI-TOF spectra of GWL-78 with sequence NF-κB-1 confirming the identity of adduct formation relating to the chromatographic peaks shown in **Figure** (NF-κB-1 observed mass: 7033.2 m/z, theoretical mass: 7032.6 m/z, GWL-78/NF-κB-1 adduct observed mass: 7624 m/z, theoretical mass: 7623.2 m/z). (**A**) NF-κB-1 alone; (**B**) after 5 min; (**C**) 3 hours; (**D**) 24 hours incubation. Refers to Chapter 2, p 106 in thesis.

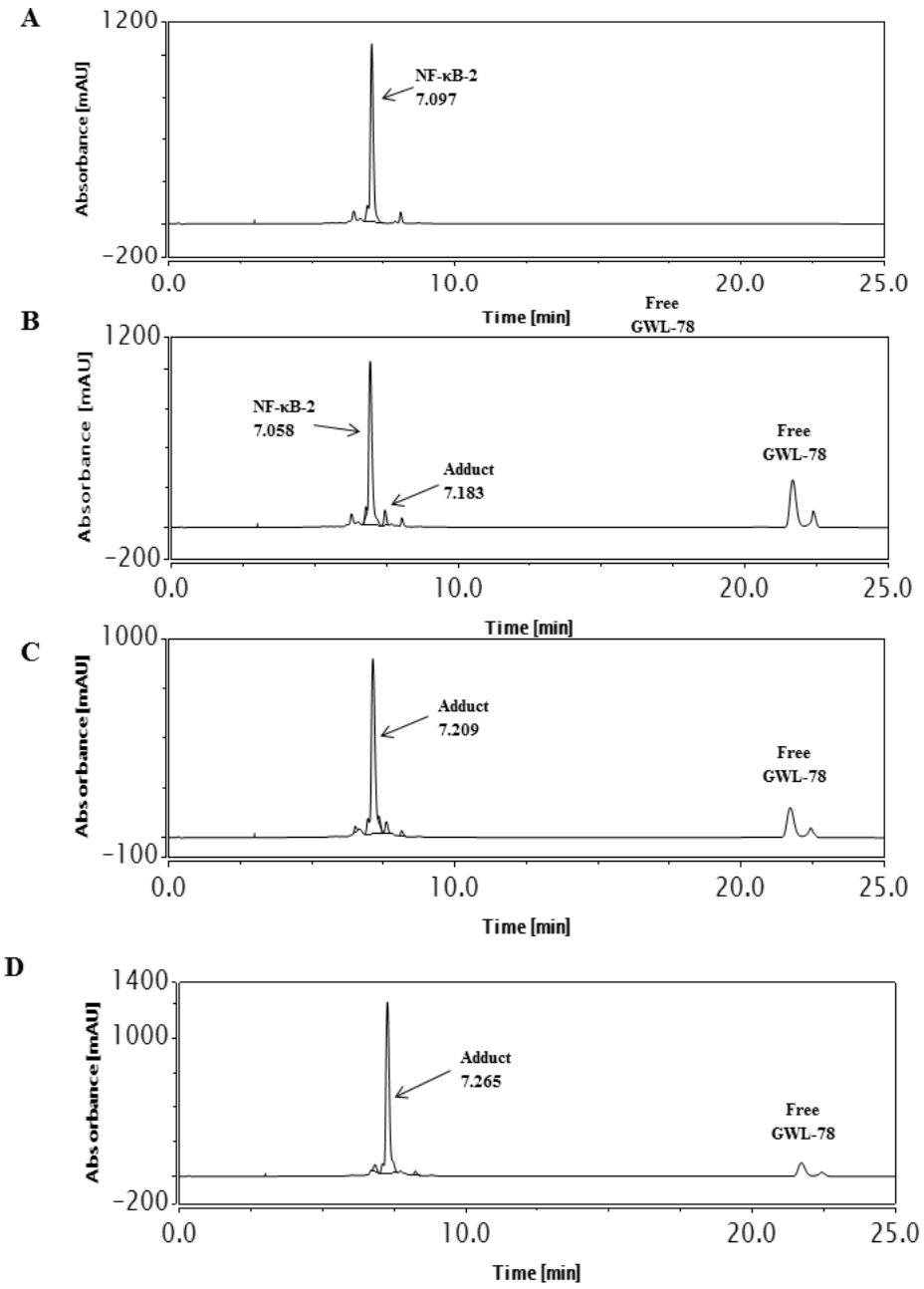


Figure A2.4: Time-course HPLC chromatograms: (**A**) Annealed NF-κB-2 sequence; (**B**) Annealed NF-κB-2 sequence after incubating with GWL-78 for 5 min showing the appearance of one new peak at RT 7.183 min (**adduct**); (**C**) 3 hours incubation; (**D**) after 24 hours incubation showing increase in adduct formation at RT 7.265 min and complete disappearance of the NF-κB-2 sequence peak at RT 7.097 min with reaction complete after 24 hours. Refers to Chapter 2, p 97 in thesis.

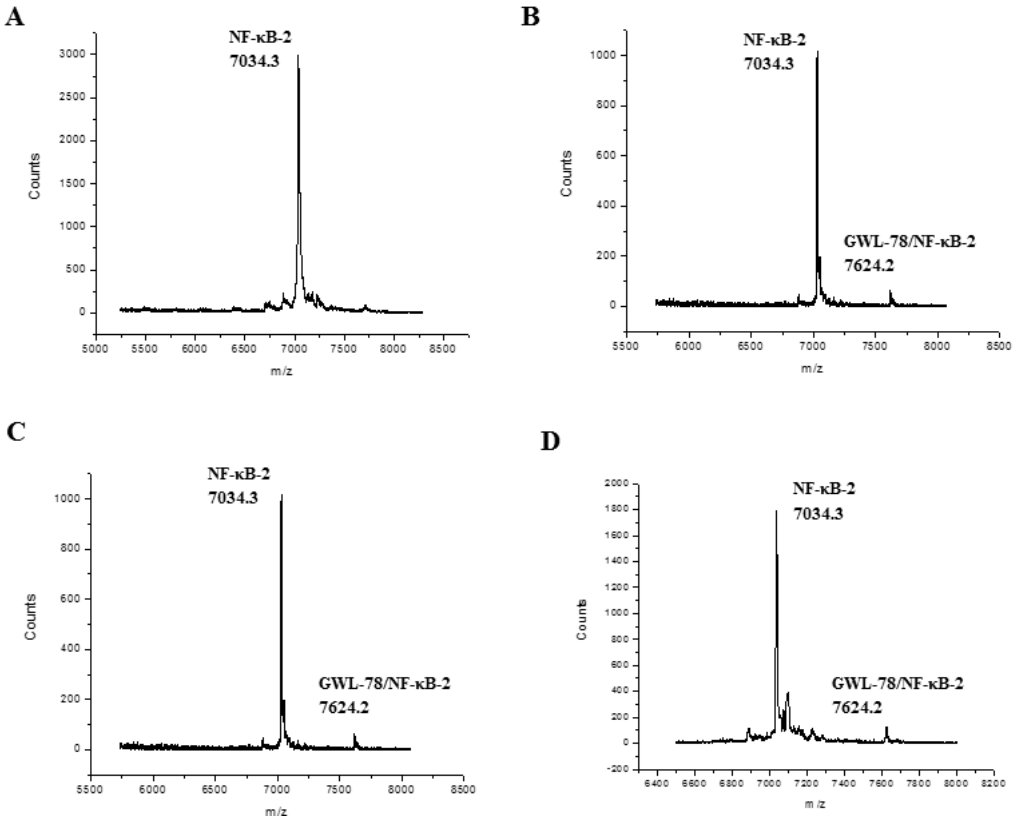


Figure A2.5: MALDI-TOF spectra of GWL-78 with sequence NF-κB-2 confirming the identity of adduct formation relating to the chromatographic peaks shown in **Figure** (NF-κB-2 observed mass: 7034.3 m/z, theoretical mass: 7032.6 m/z, GWL-78/NF-κB-2 adduct observed mass: 7625.1 m/z, theoretical mass: 7623.2 m/z). (**A**) NF-κB-2 alone; (**B**) after 5 min; (**C**) 3 hours; (**D**) 24 hours incubation. Refers to Chapter 2, p 106 in thesis.

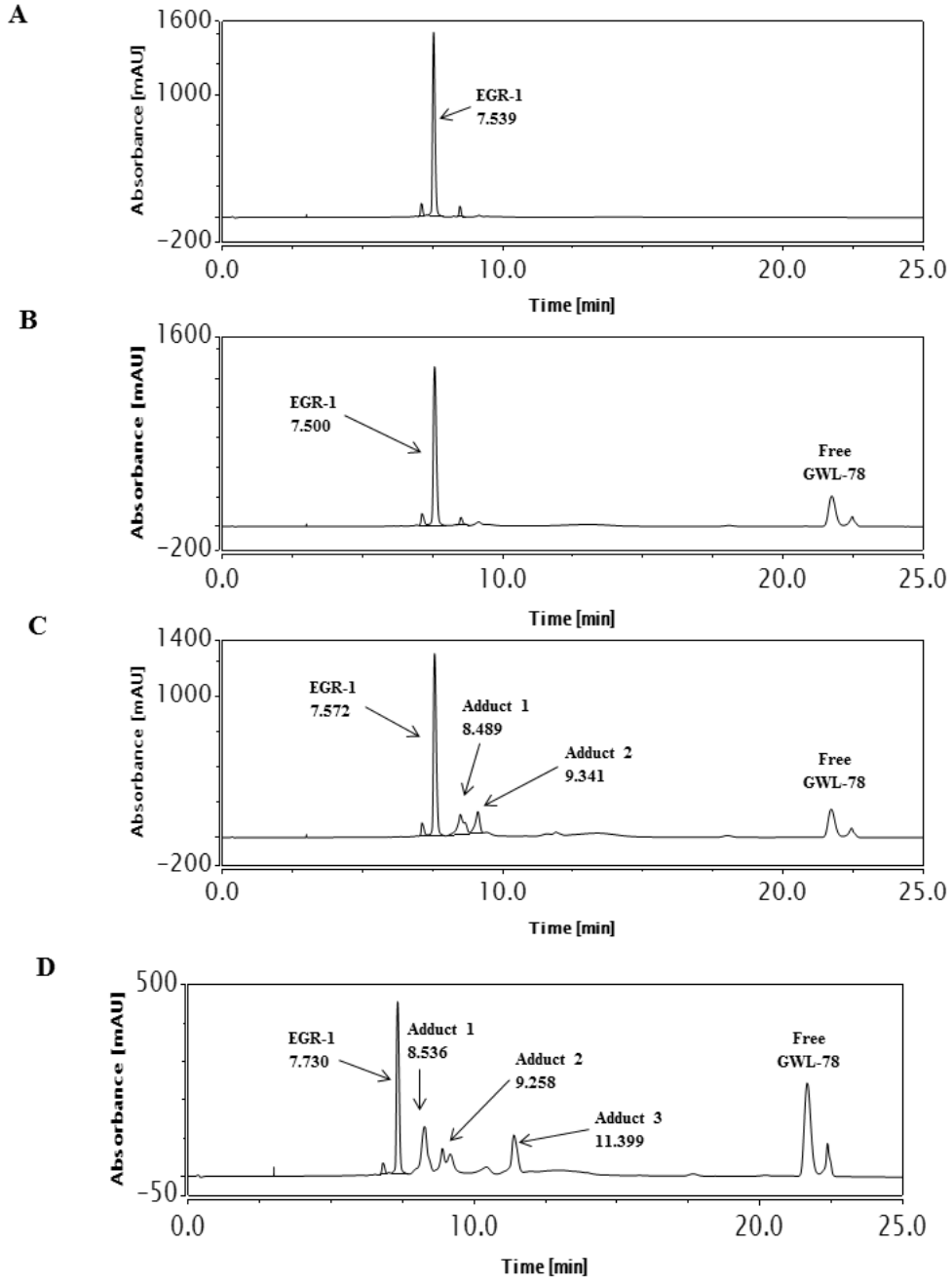


Figure A2.6: Time-course HPLC chromatograms: (**A**) Annealed EGR-1 sequence; (**B**) Annealed EGR-1 sequence after incubating with GWL-78 for 5 min showing no changes in the chromatogram; (**C**) after 3 hours; (**D**) after 24 hours incubation showing the appearance of three different adducts at RT 8.536 min (**adduct 1**), 9.258 min (**adduct 2**), and 11.399 min (**adduct 3**) with reaction not complete after 24 hours. Refers to Chapter 2, p 97 in thesis.

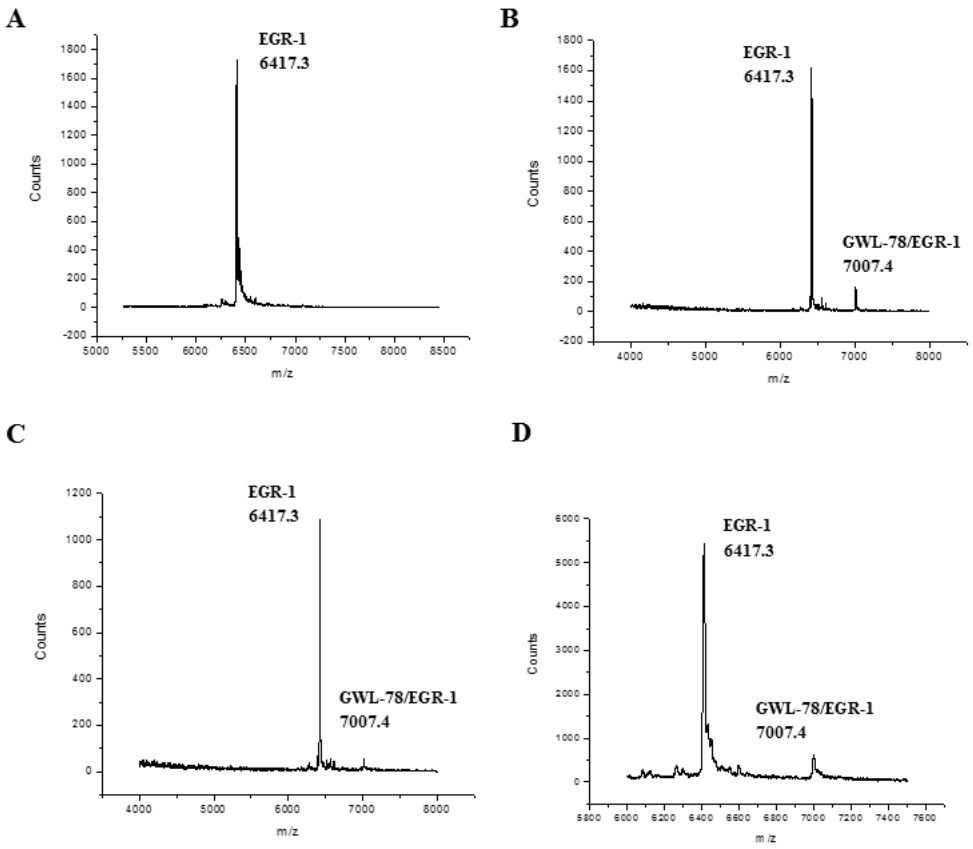


Figure A2.7: MALDI-TOF spectra of GWL-78 with sequence EGR-1 confirming the identity of adduct formation relating to the chromatographic peaks shown in **Figure** (EGR-1 observed mass: 6417.3 m/z, theoretical mass: 6416.2 m/z, GWL-78/EGR-1 adduct observed mass: 7007.4 m/z, theoretical mass: 7006.8 m/z). (**A**) EGR-1 alone; (**B**) after 5 min; (**C**) 3 hours; (**D**) 24 hours incubation. Refers to Chapter 2, p 106 in thesis.

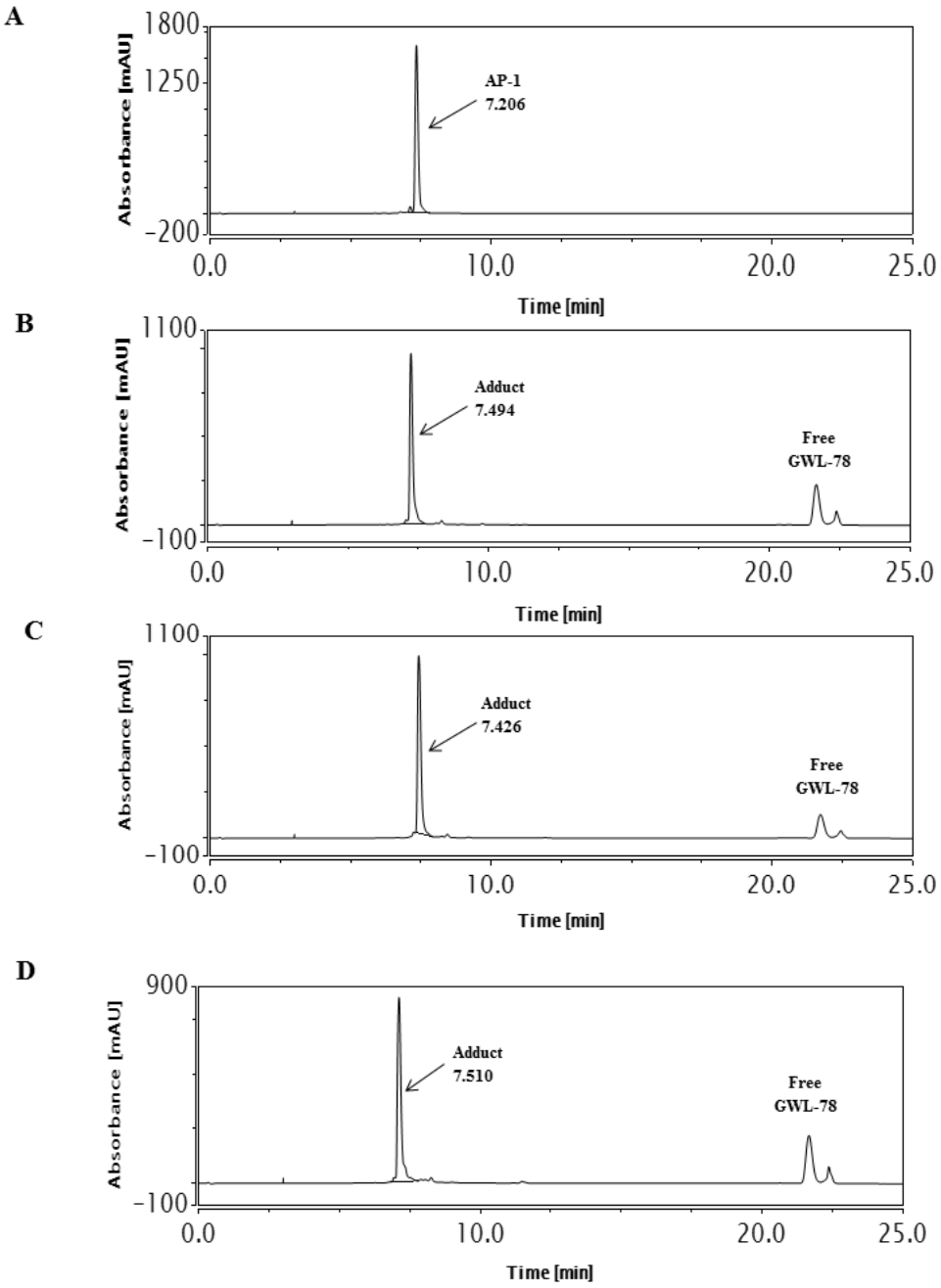


Figure A2.8: Time-course HPLC chromatograms: (**A**) Annealed AP-1 sequence; (**B**) Annealed AP-1 sequence after incubating with GWL-78 for 5 min showing the appearance of a new major peak at RT 7.494 (**adduct**) min with disappearance of the DNA peak indicating reaction is complete; (**C**) after 3 hours; (**D**) after 24 hours incubation showing no further changes in the chromatograms with the adduct peak at RT 7.510 min still remaining. Refers to Chapter 2, p 97 in thesis.

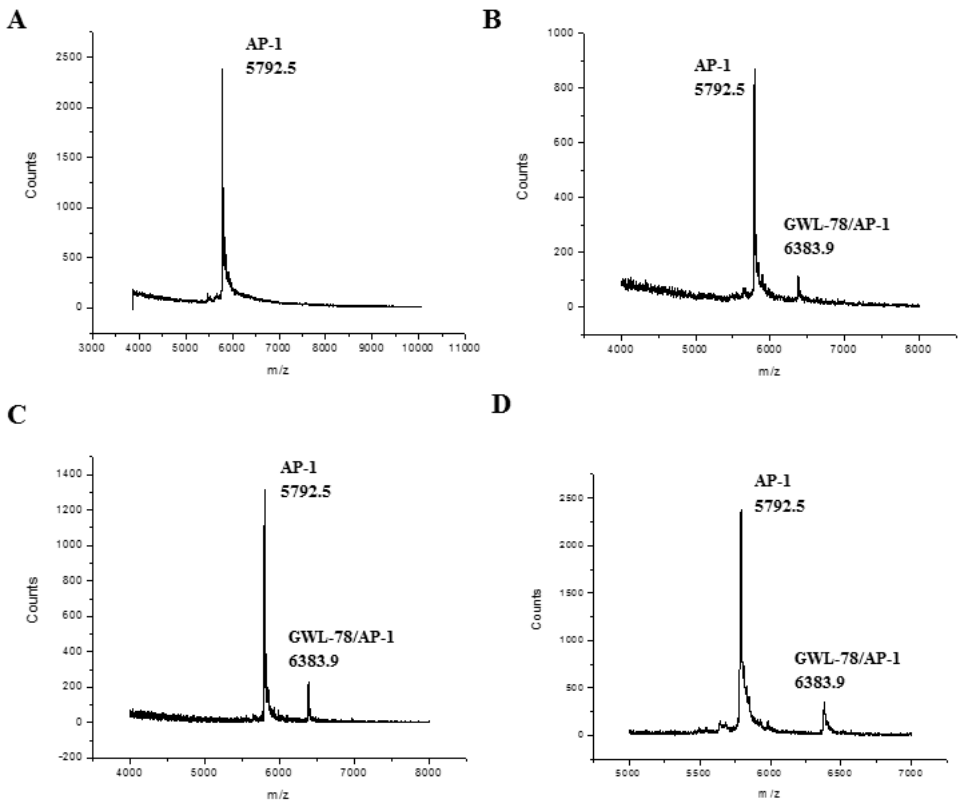


Figure A2.9: MALDI-TOF spectra of GWL-78 with sequence AP-1 confirming the identity of adduct formation relating to the chromatographic peaks shown in **Figure** (AP-1 observed mass: 5792.5 m/z, theoretical mass: 5793.8 m/z, GWL-78/AP-1 adduct observed mass: 6383.9 m/z, theoretical mass 6384.4 m/z). (**A**) AP-1 alone; (**B**) after 5 min; (**C**) 3 hours; (**D**) 24 hours incubation. Refers to Chapter 2, p 106 in thesis.

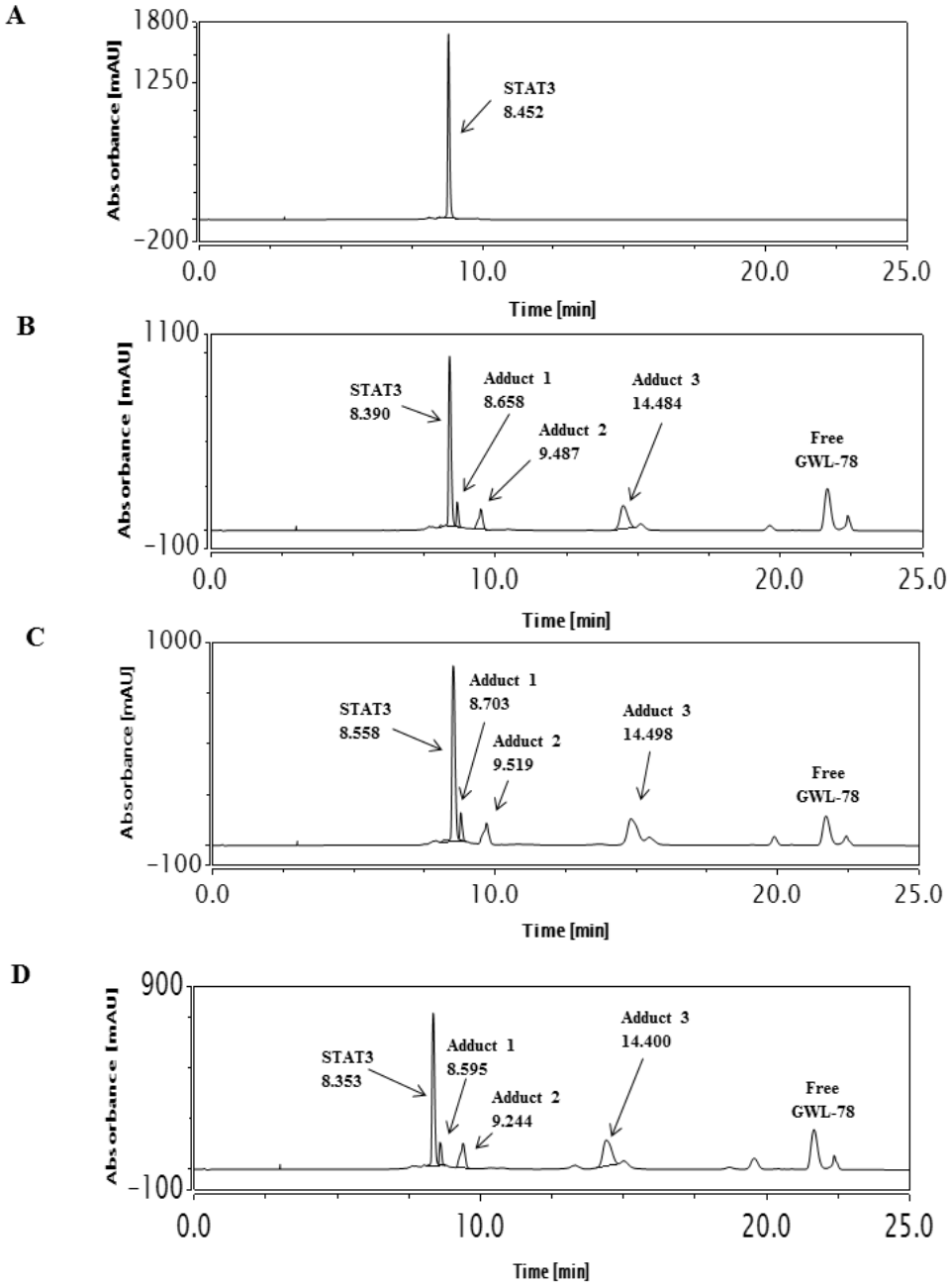


Figure A2.10: Time-course HPLC chromatograms: (**A**) Annealed STAT3 sequence; (**B**) Annealed STAT3 sequence after incubating with GWL-78 for 5 min showing the appearance of three new minor peaks at RT 8.658 min (**adduct 1**), RT 9.487 min (**adduct 2**) and 14.484 min (**adduct 3**); (**C**) after 3 hours; (**D**) after 24 hours incubation showing no further changes in the chromatograms with the adduct peaks at RT 8.595 min, RT 9.244 min and RT 14.400 min still remaining indicating reaction is not complete after 24 hours. Refers to Chapter 2, p 97 in thesis.

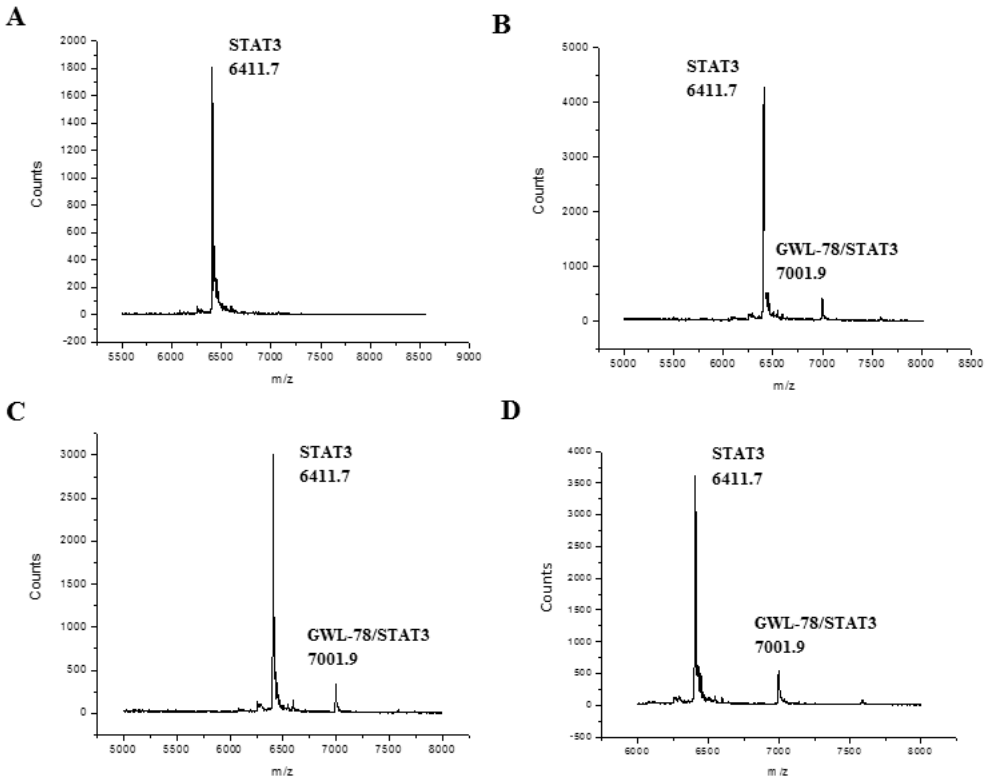


Figure A2.11: MALDI-TOF spectra of GWL-78 with sequence STAT3 confirming the identity of adduct formation relating to the chromatographic peaks shown in **Figure** (STAT3 observed mass: 6411.7 m/z, theoretical mass: 6412.2 m/z, GWL-78/STAT3 adduct observed mass: 7001.9 m/z, theoretical mass 7002.8 m/z). (**A**) STAT3 alone; (**B**) after 5 min; (**C**) 3 hours; (**D**) 24 hours incubation. Refers to Chapter 2, p 106 in thesis.

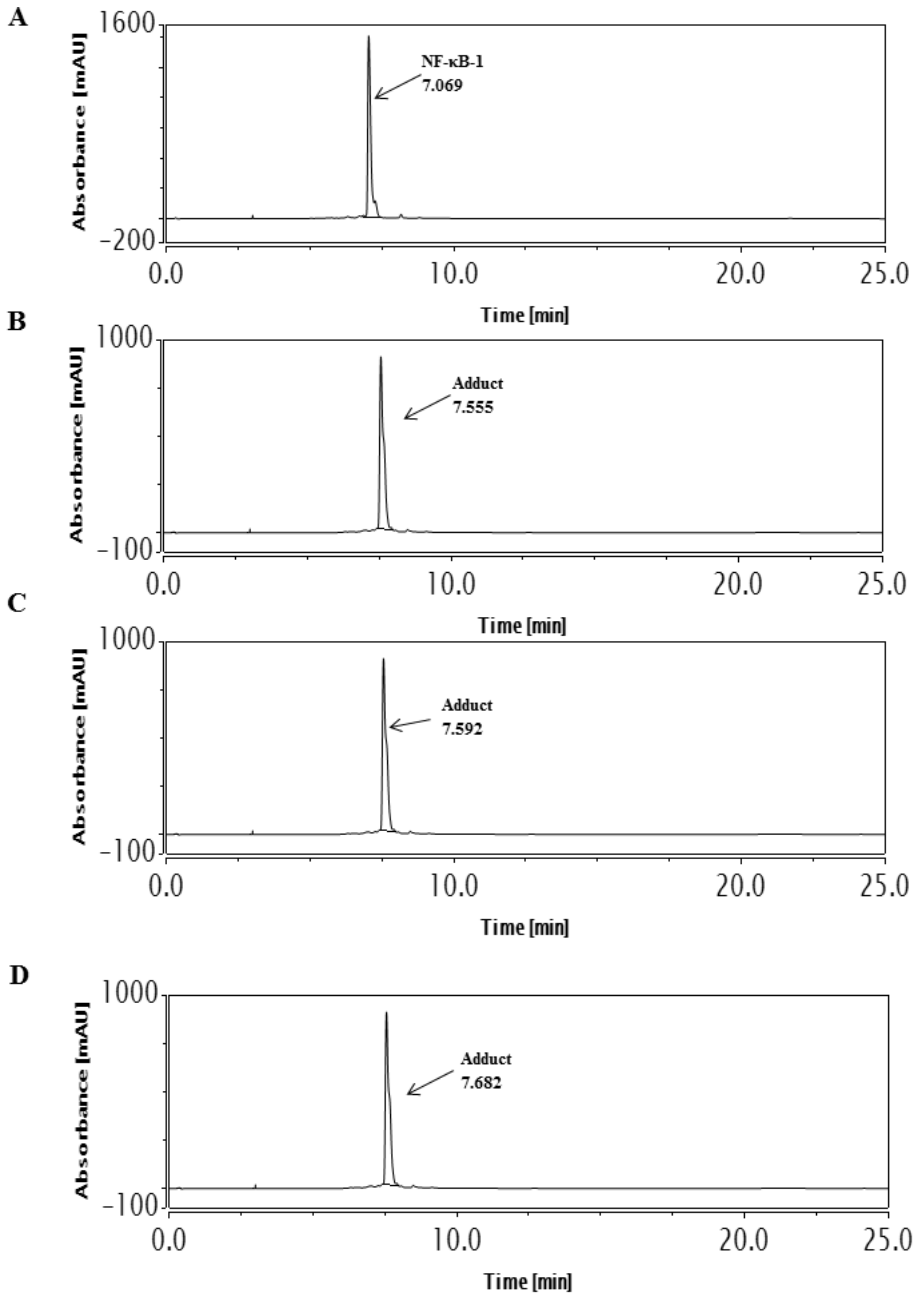


Figure A2.12: Time-course HPLC chromatograms: (**A**) Annealed NF-κB-1 sequence; (**B**) Annealed NF-κB-1 sequence after incubating with KMR-28-39 for 5 min showing the appearance of a new major peak at RT 7.555 min (**adduct**) with disappearance of the DNA peak indicating reaction is complete; (**C**) after 3 hours; (**D**) after 24 hours incubation showing no further changes in the chromatograms with the adduct peak at RT 7.682 min still remaining. Refers to Chapter 2, p 101 in thesis.

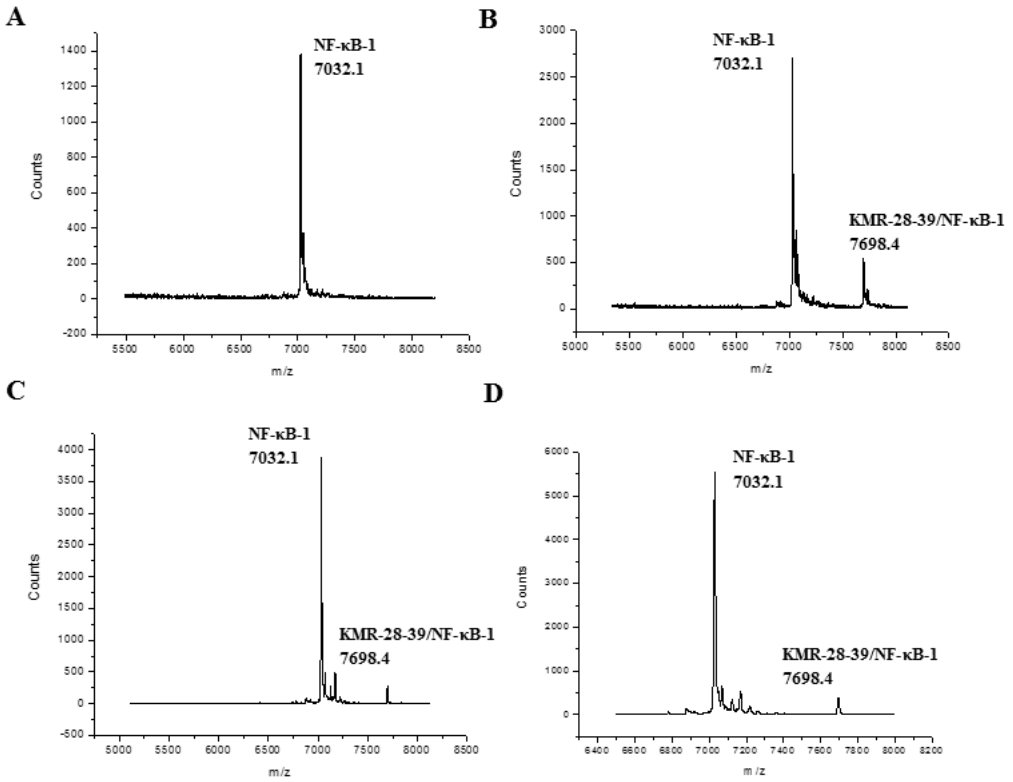


Figure A2.13: MALDI-TOF spectra of KMR-28-39 with sequence NF-κB-1 confirming the identity of adduct formation relating to the chromatographic peaks shown in **Figure** (NF-κB-1 observed mass: 7032.1 m/z, theoretical mass: 7032.6 m/z, KMR-28-39/NF-κB-1 adduct observed mass: 7698.4 m/z, theoretical mass: 7699.6 m/z). (**A**) NF-κB-1 alone; (**B**) after 5 min; (**C**) 3 hours; (**D**) 24 hours incubation. Refers to Chapter 2, p 106 in thesis.

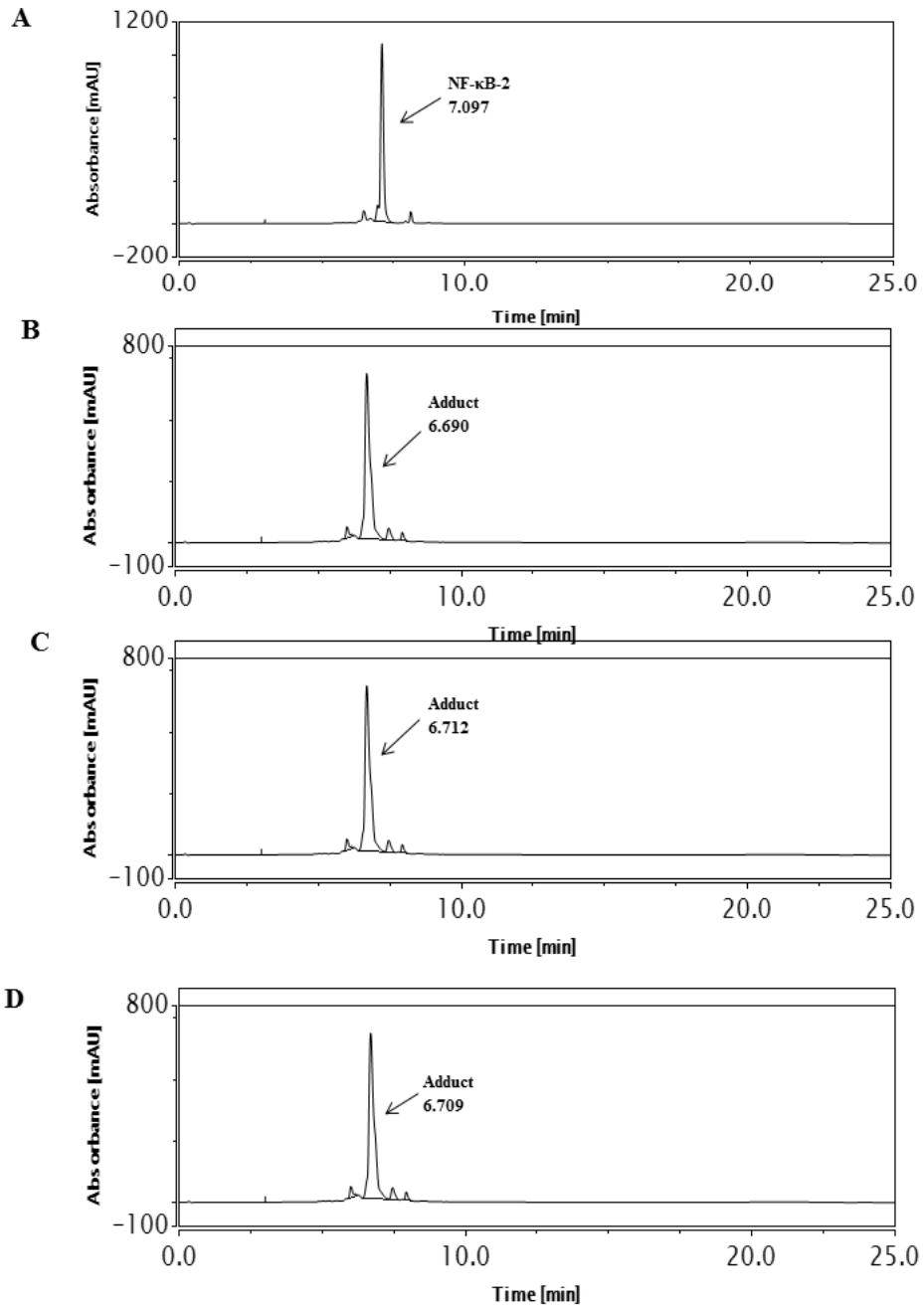


Figure A2.14: Time-course HPLC chromatograms: (**A**) Annealed NF-κB-2 sequence; (**B**) Annealed NF-κB-2 sequence after incubating with KMR-28-39 for 5 min showing the appearance of a new major peak at RT 6.690 min (**adduct**) with disappearance of the DNA peak indicating reaction is complete; (**C**) after 3 hours; (**D**) after 24 hours incubation showing no further changes in the chromatograms with the adduct peak at RT 6.709 min still remaining. Refers to Chapter 2, p 101 in thesis.

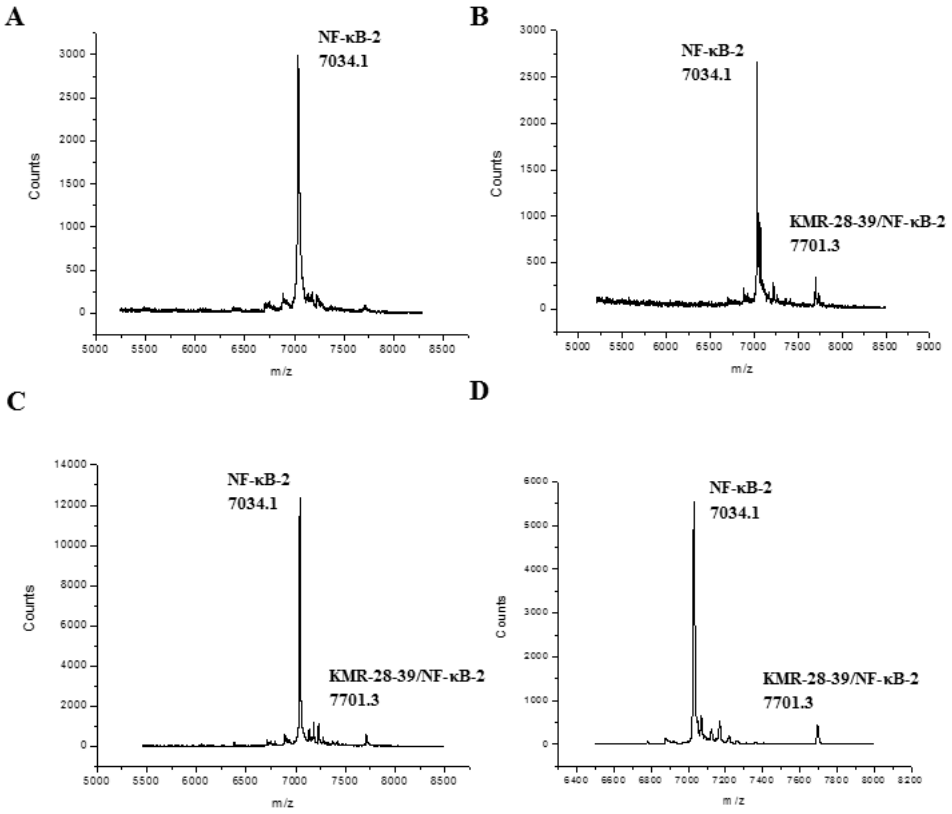


Figure A2.15: MALDI-TOF spectra of KMR-28-39 with sequence NF-κB-2 confirming the identity of adduct formation relating to the chromatographic peaks shown in **Figure** (NF-κB-2 observed mass: 7034.1 m/z, theoretical mass: 7033.6 m/z, KMR-28-39/NF-κB-2 adduct observed mass: 7701.3 m/z, theoretical mass 7700.6 m/z). (**A**) NF-κB-2 alone; (**B**) after 5 min; (**C**) 3 hours; (**D**) 24 hours incubation. Refers to Chapter 2, p 106 in thesis.

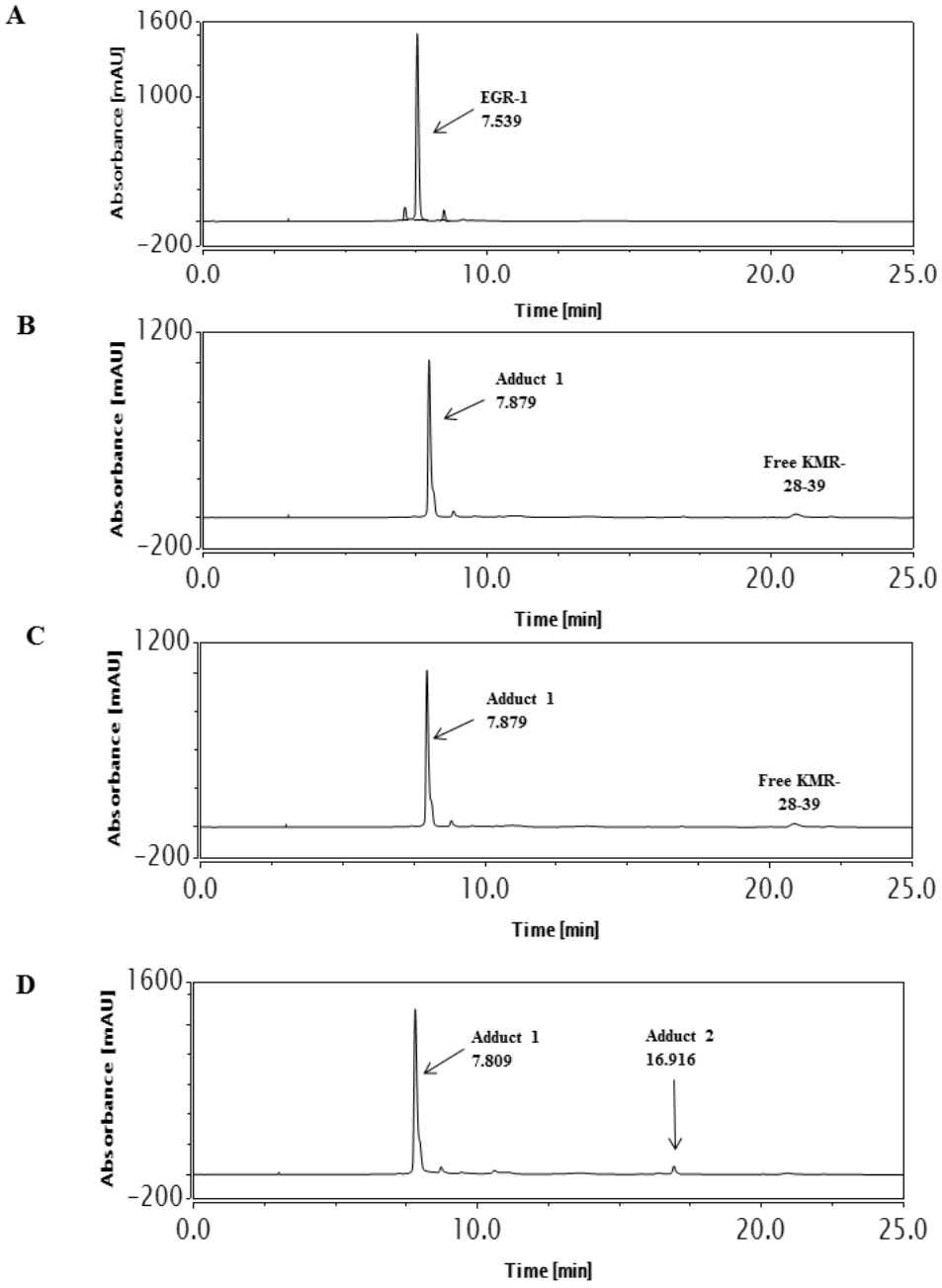


Figure A2.16: Time-course HPLC chromatograms: (**A**) Annealed EGR-1 sequence; (**B**) Annealed EGR-1 sequence after incubating with KMR-28-39 for 5 min showing the appearance of a new major peak at RT 7.879 (**adduct**) min with disappearance of the DNA peak; (**C**) after 3 hours; (**D**) 24 hours incubation showing an new minor peak at RT 16.916 min (**adduct 2**). Refers to Chapter 2, p 101 in thesis.

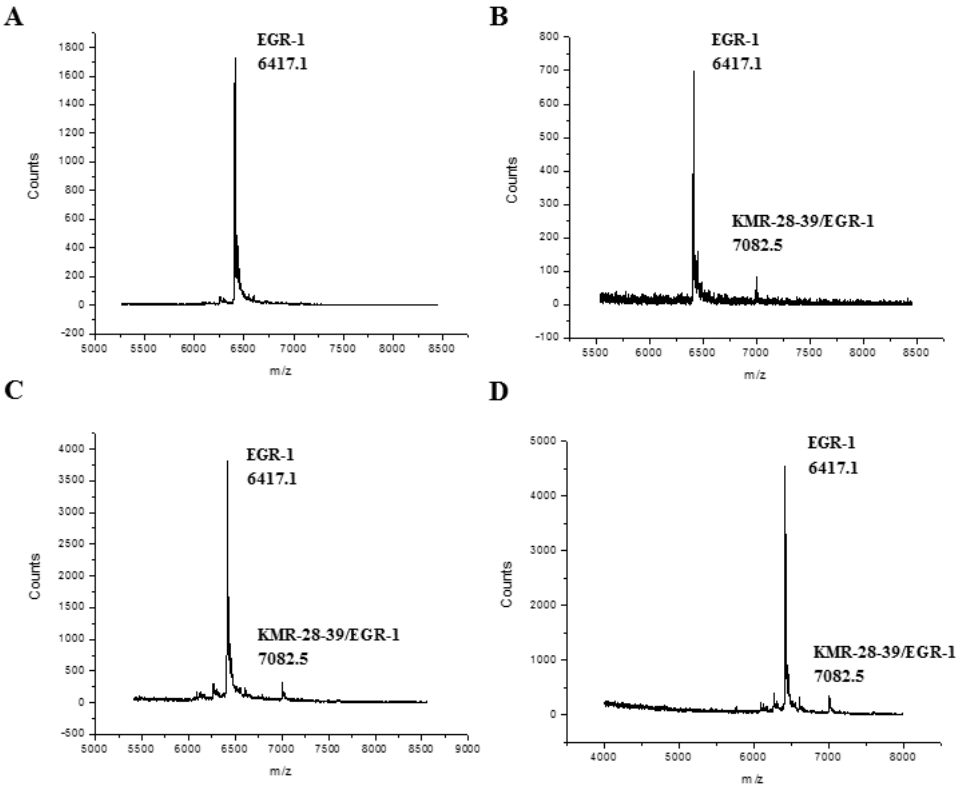


Figure A2.17: MALDI-TOF spectra of KMR-28-39 with sequence EGR-1 confirming the identity of adduct formation relating to the chromatographic peaks shown in **Figure** (EGR-1 observed mass: 6417.1 m/z, theoretical mass: 6416.2 m/z, KMR-28-39/EGR-1 adduct observed mass: 7082.5 m/z, theoretical mass 7083.2 m/z). (**A**) EGR-1 alone; (**B**) after 5 min; (**C**) 3 hours; (**D**) 24 hours incubation. Refers to Chapter 2, p 106 in thesis.

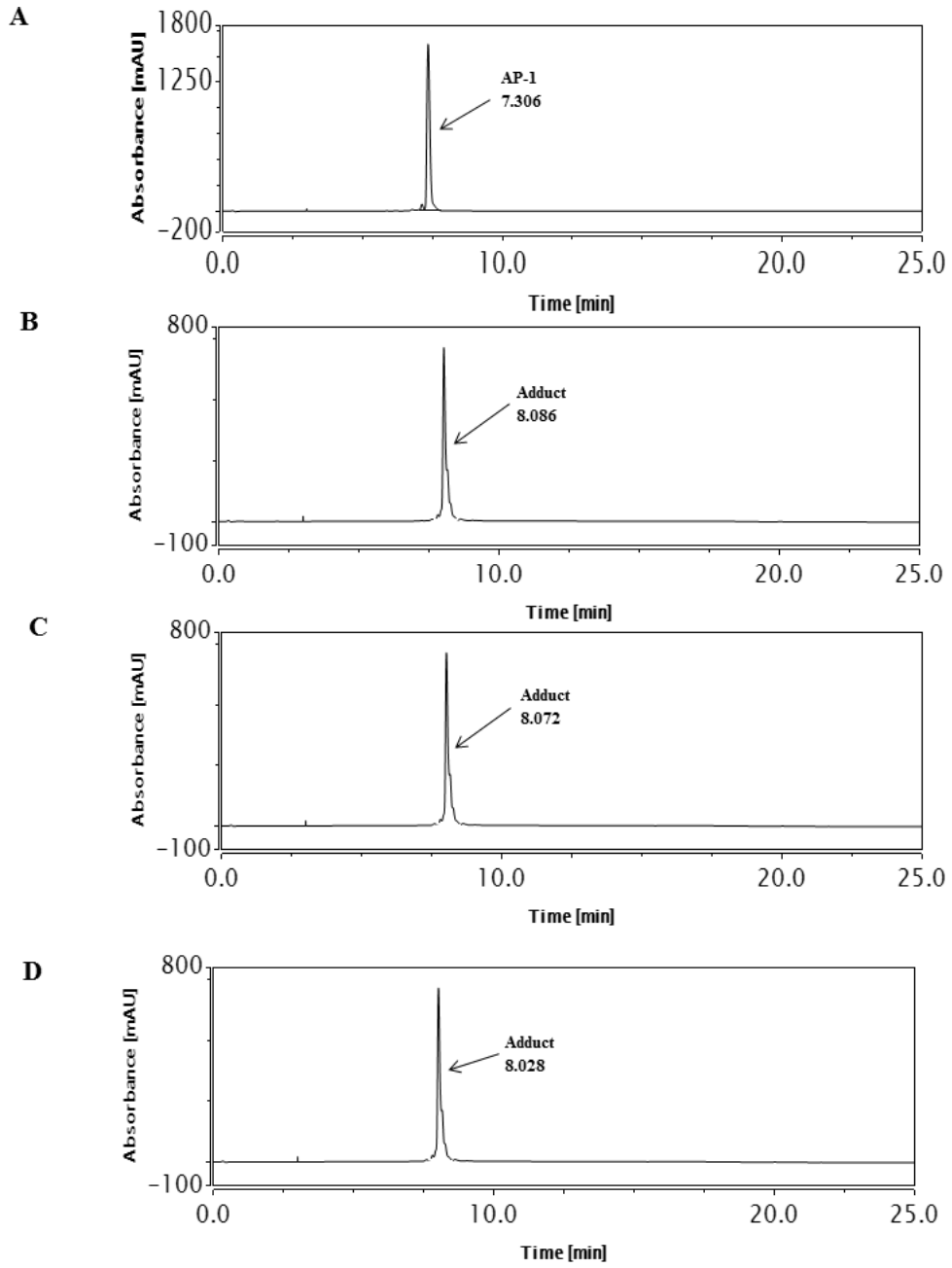


Figure A2.18: Time-course HPLC chromatograms: (**A**) Annealed AP-1 sequence; (**B**) Annealed AP-1 sequence after incubating with KMR-28-39 for 5 min showing the appearance of a new major peak at RT 8.086 min (**adduct**) with disappearance of the DNA peak indicating reaction is complete; (**C**) after 3 hours; (**D**) after 24 hours incubation showing no further changes in the chromatograms with the adduct peak at RT 8.028 min still remaining. Refers to Chapter 2, p 101 in thesis.

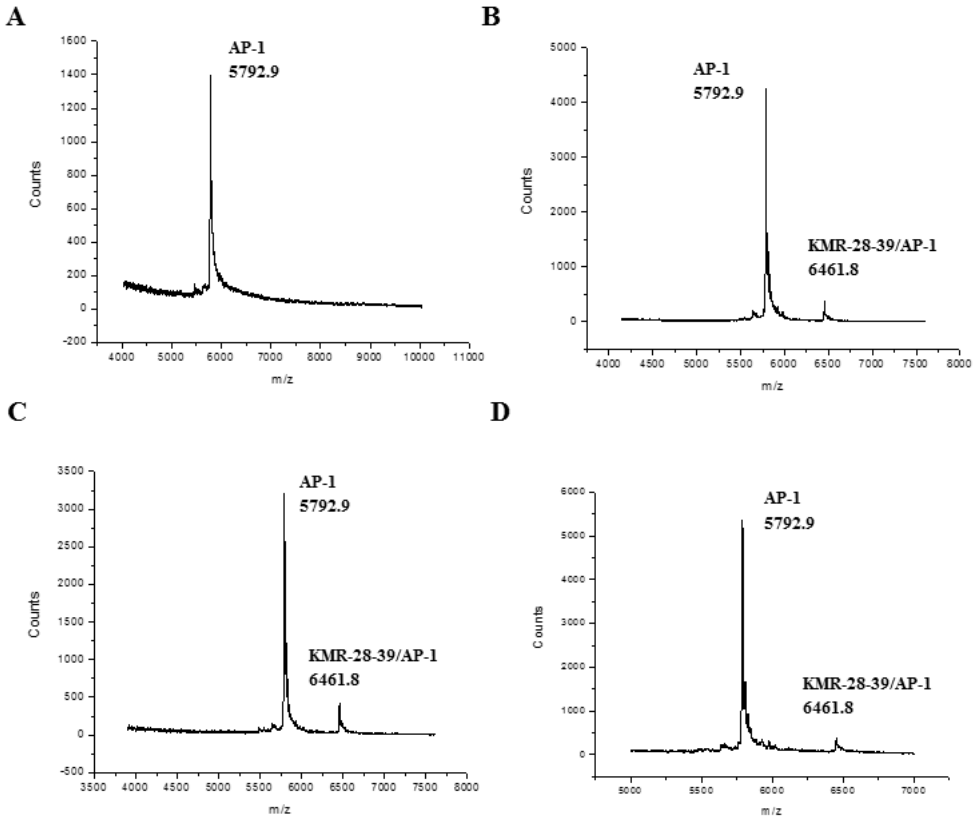


Figure A2.19: MALDI-TOF spectra of KMR-28-39 with sequence AP-1 confirming the identity of adduct formation relating to the chromatographic peaks shown in **Figure** (AP-1 observed mass: 5792.9 m/z, theoretical mass: 5793.8 m/z, KMR-28-39/AP-1 adduct observed mass: 6461.8 m/z, theoretical mass 6460.8 m/z). (**A**) AP-1 alone; (**B**) after 5 min; (**C**) 3 hours; (**D**) 24 hours incubation. Refers to Chapter 2, p 106 in thesis.

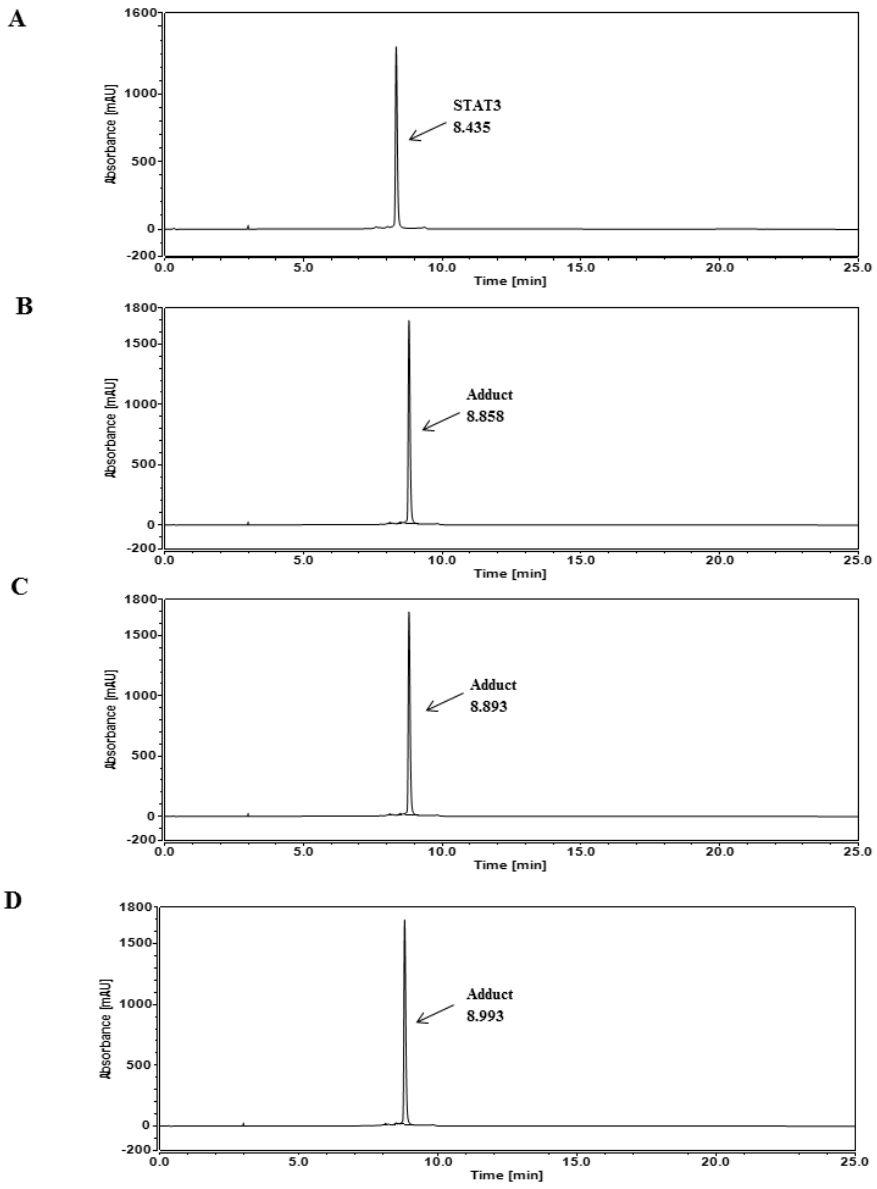


Figure A2.20: Time-course HPLC chromatograms: (**A**) Annealed STAT3 sequence; (**B**) Annealed STAT3 sequence after incubating with KMR-28-39 for 5 min showing the appearance of a new major peak at RT 8.858 min (**adduct**) with disappearance of the DNA peak indicating reaction is complete; (**C**) after 3 hours; (**D**) after 24 hours incubation showing no further changes in the chromatograms with the adduct peak at RT 8.993 min still remaining. Refers to Chapter 2, p 101 in thesis.

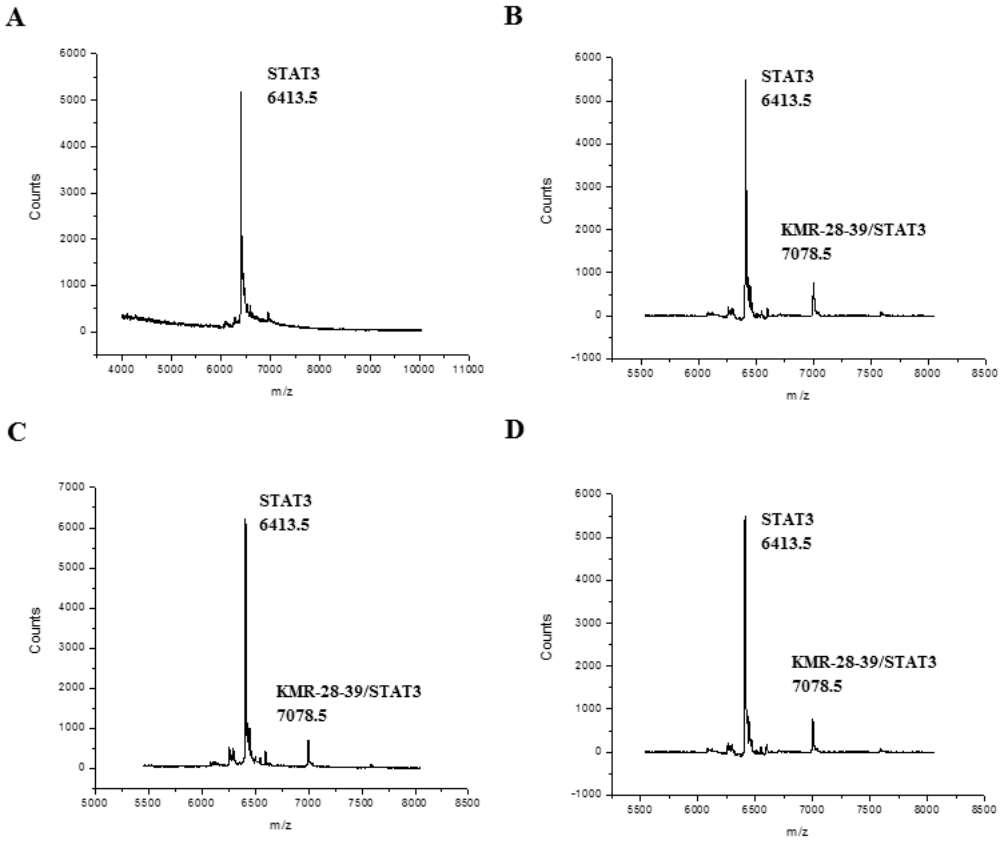


Figure A2.21: MALDI-TOF spectra of KMR-28-39 with sequence STAT3 confirming the identity of adduct formation relating to the chromatographic peaks shown in **Figure** (STAT3 observed mass: 6413.5 m/z, theoretical mass 6412.2 m/z, KMR-28-39/STAT3 adduct observed mass: 7078.5 m/z, theoretical mass 7079.2 m/z). (**A**) STAT3 alone; (**B**) after 5 min; (**C**) 3 hours; (**D**) 24 hours incubation. Refers to Chapter 2, p 106 in thesis.

Appendix Chapter 3: Additional HPLC chromatograms, MALDI-TOF spectra, FRET tables and melting curves

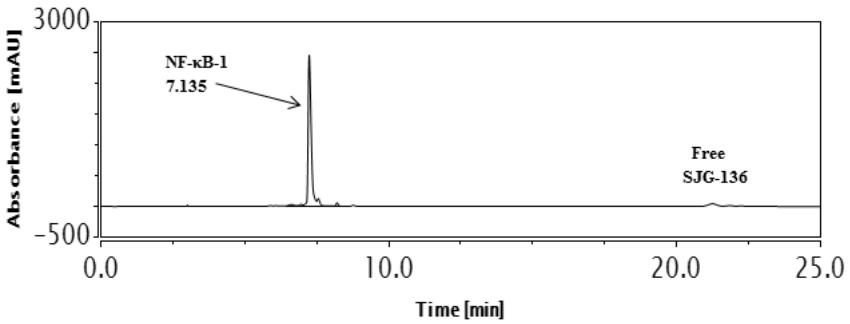


Figure A3.1: Annealed NF-κB-1 sequence after incubating with SJG-136 for 5 min showing no changes in the chromatographic profile with the DNA peak remaining at 7.135 min. Refers to Chapter 3, p 112 in thesis.

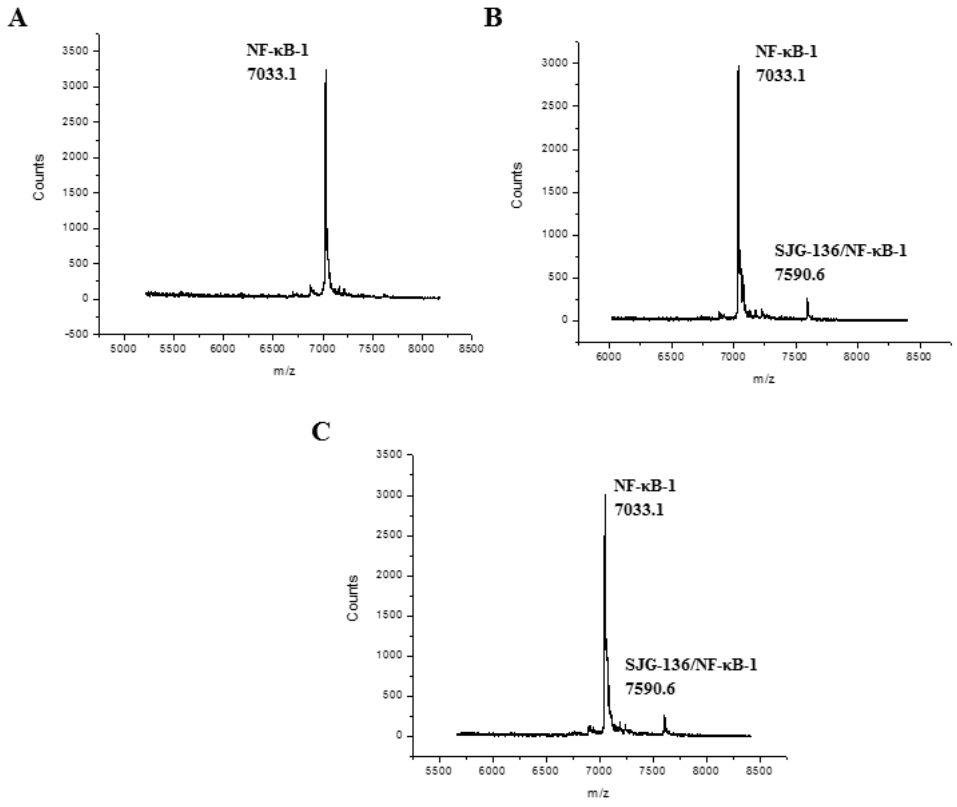


Figure A3.2: MALDI-TOF spectra of SJG-136 with sequence NF-κB-1 confirming the identity of adduct formation relating to the chromatographic peaks shown in **Figure 3.2** (NF-κB-1 observed mass: 7033.1 m/z, theoretical mass: 7032.6 m/z, SJG-136/NF-κB-1 adduct observed mass: 7590.6 m/z, theoretical mass: 7589.2 m/z). (**A**) NF-κB-1 alone; (**B**) after 5 min; (**C**) 3 hours incubation. Refers to Chapter 3, p 112 in thesis.

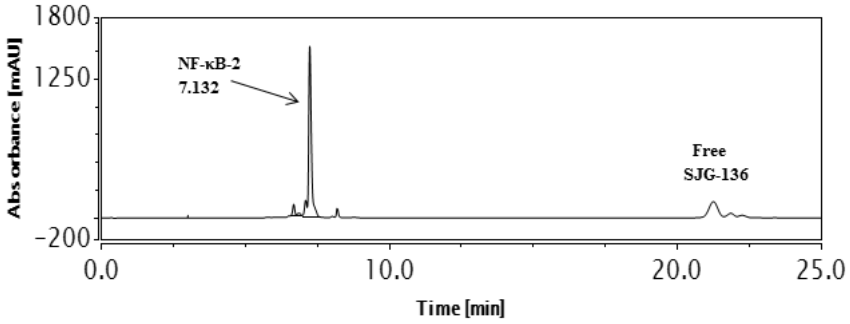


Figure A3.3: Annealed NF-κB-2 sequence after incubating with SJG-136 for 5 min showing no changes in the chromatographic profile with the DNA peak remaining at 7.132 min. Refers to Chapter 3, p 114 in thesis.

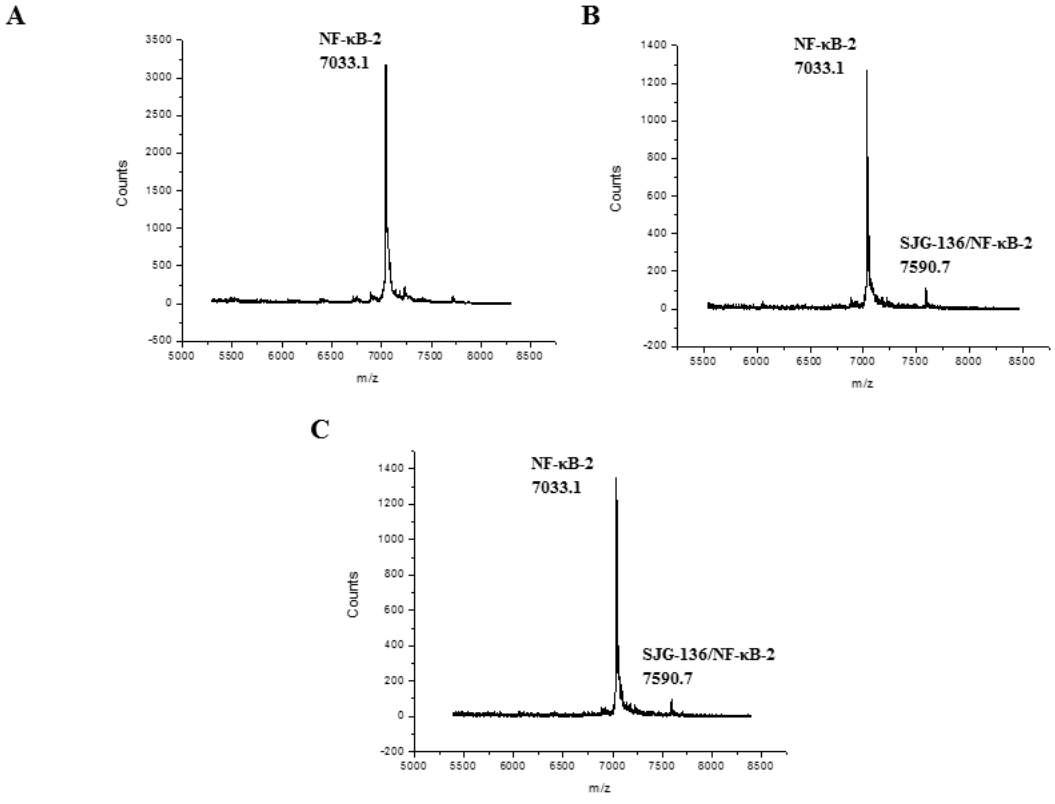


Figure A3.4: MALDI-TOF spectra of SJG-136 with sequence NF-κB-2 confirming the identity of adduct formation relating to the chromatographic peaks shown in **Figure 3.3** (NF-κB-2 observed mass: 7033.1 m/z, theoretical mass: 7033.6 m/z, SJG-136/NF-κB-2 adduct observed mass: 7590.7 m/z, theoretical mass: 7590.2 m/z). (**A**) NF-κB-2 alone; (**B**) after 5 min; (**C**) 3 hours incubation. Refers to Chapter 3, p 114 in thesis.

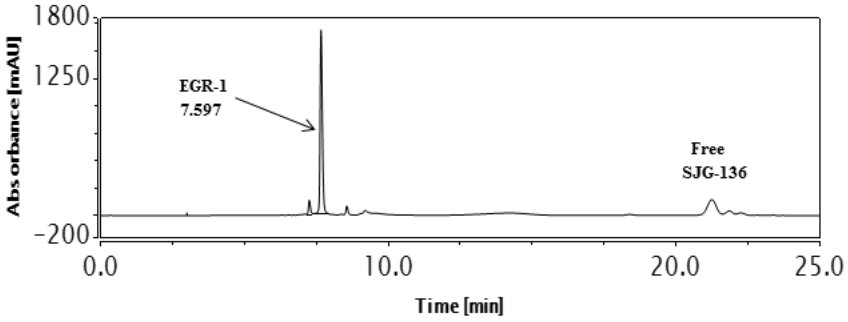


Figure A3.5: Annealed EGR-1 sequence after incubating with SJG-136 for 5 mon showing no changes in the chromatographic profile with the DNA peak remaining at 7.597 min. Refers to Chapter 3, p 116 in thesis.

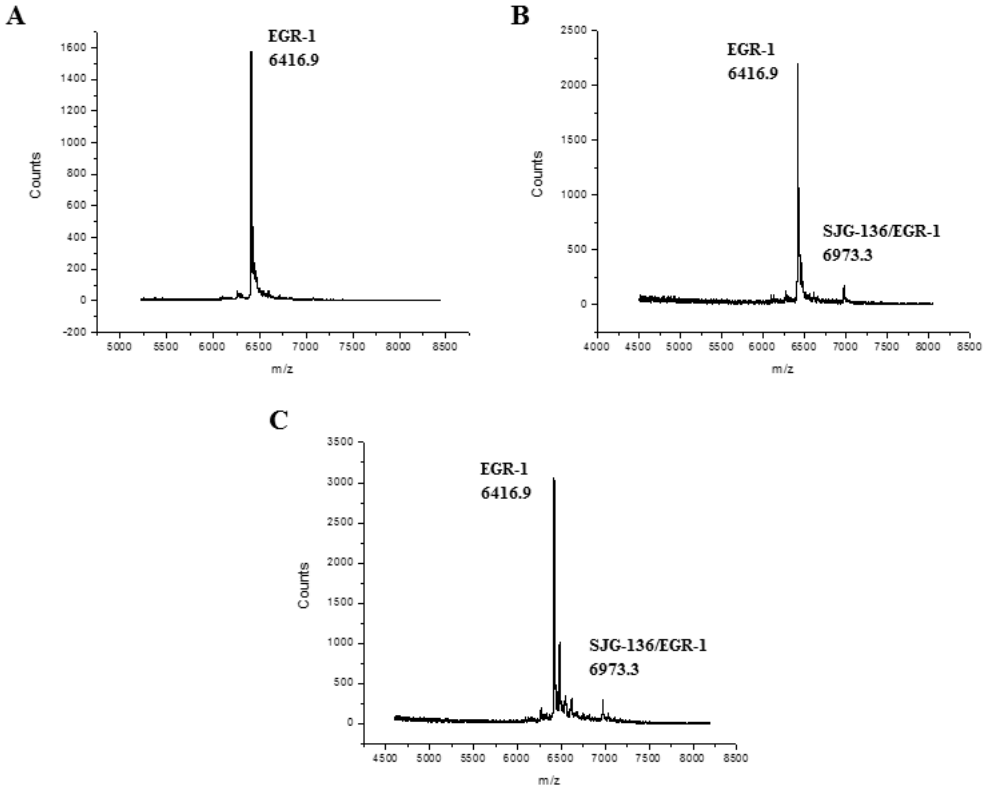


Figure A3.6: MALDI-TOF spectra of SJG-136 with sequence EGR-1 confirming the identity of adduct formation relating to the chromatographic peaks shown in **Figure 3.4** (EGR-1 observed mass: 6416.9 m/z, theoretical mass: 6416.2 m/z, SJG-136/EGR-1 adduct observed mass: 6973.3 m/z, theoretical mass: 6972.8 m/z). (**A**) EGR-1 alone; (**B**) after 5 min; (**C**) 3 hours incubation. Refers to Chapter 3, p 116 in thesis.

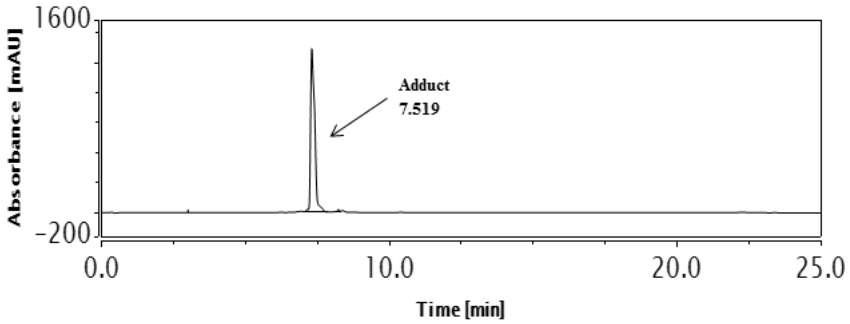


Figure A3.7: Annealed AP-1 sequence after incubating with SJG-136 for 5 min showing the appearance of a new major peak at 7.519 min. Refers to Chapter 3, p 118 in thesis.

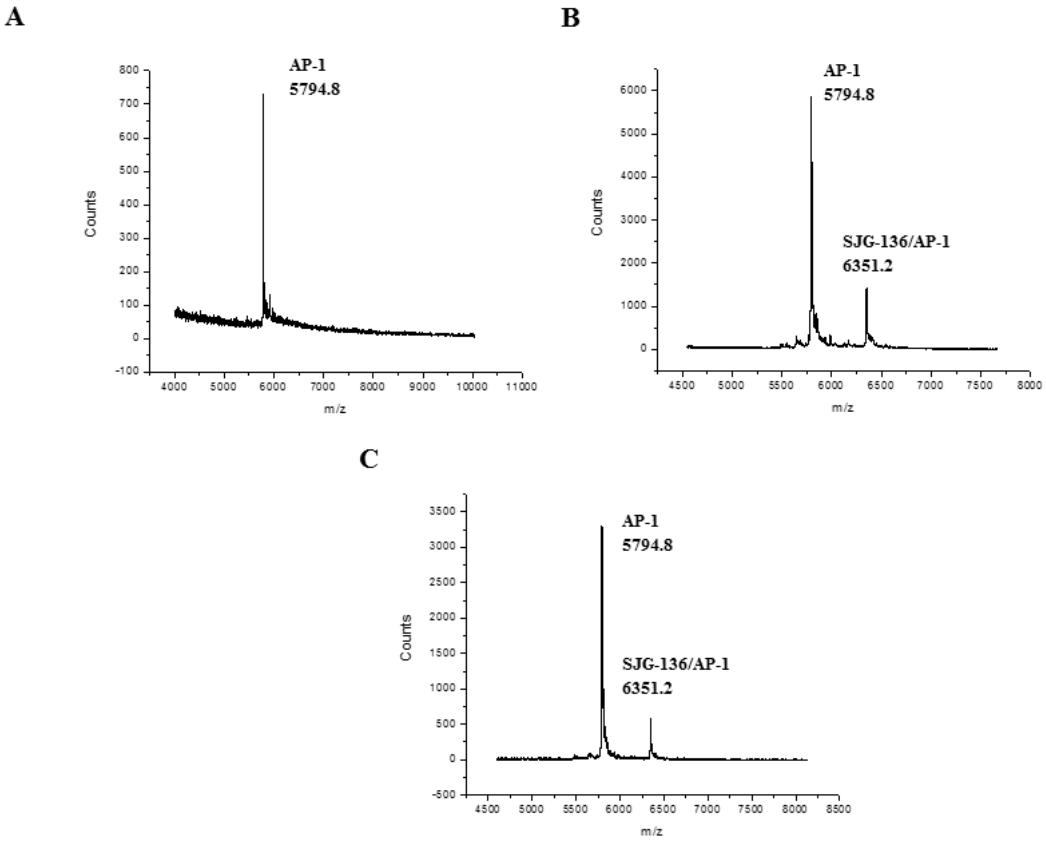


Figure A3.8: MALDI-TOF spectra of SJG-136 with sequence AP-1 confirming the identity of adduct formation relating to the chromatographic peaks shown in **Figure 3.5** (AP-1 observed mass: 5794.8 m/z, theoretical mass: 5793.8 m/z, SJG-136/AP-1 adduct observed mass: 6351.2 m/z, theoretical mass: 6350.4 m/z). (**A**) AP-1 alone; (**B**) after 5 min; (**C**) 3 hours incubation. Refers to Chapter 3, p 118 in thesis.

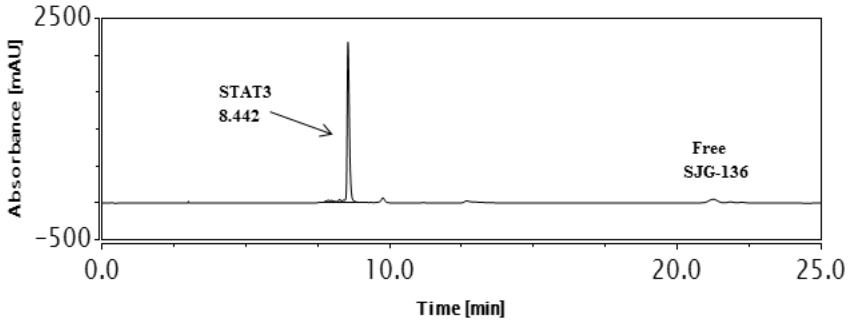


Figure A3.9: Annealed STAT3 sequence after incubating with SJG-136 for 5 min showing no changes in the chromatographic profile with the DNA peak remaining at 8.442 min. Refers to Chapter 3, p 120 in thesis.

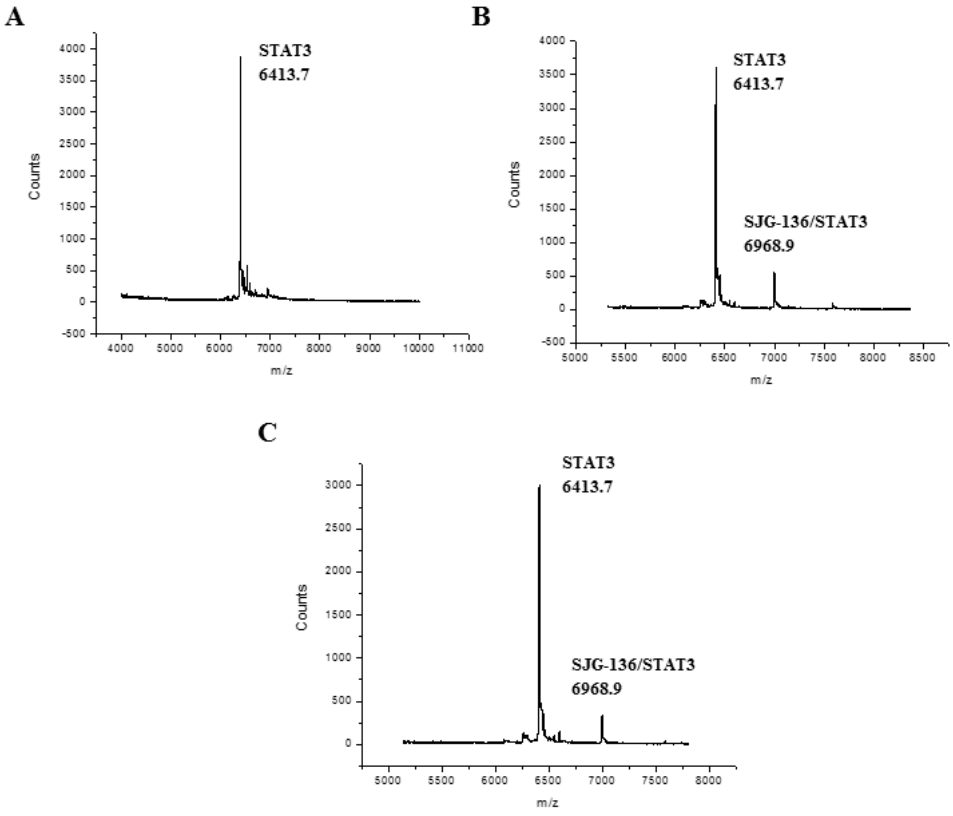


Figure A3.10: MALDI-TOF spectra of SJG-136 with sequence STAT3 confirming the identity of adduct formation relating to the chromatographic peaks shown in **Figure 3.6** (STAT3 observed mass: 6413.7 m/z, theoretical mass: 6412.2 m/z, SJG-136/STAT3 adduct observed mass: 6968.9 m/z, theoretical mass: 6968.8 m/z). (**A**) STAT3 alone; (**B**) after 5 min; (**C**) 3 hours incubation. Refers to Chapter 3, p 120 in thesis.

Table A3.1: Average melting temperature [°C] of the hairpin sequences after treatment with SJG-136 at various concentrations (FRET). Refers to Chapter 3, p 129 in thesis.

SJG-136 [μM]	NF-κB-1	NF-κB-2	EGR-1	AP-1	STAT3
100	96.1	95.0	93.7	87.8	65.4
50	96.1	94.7	89.8	88.1	64.0
10	79.3	85.9	84.2	88.1	62.4
1	79.5	84.4	83.4	68.5	61.6
0.1	78.9	84.5	84.1	68.2	61.6

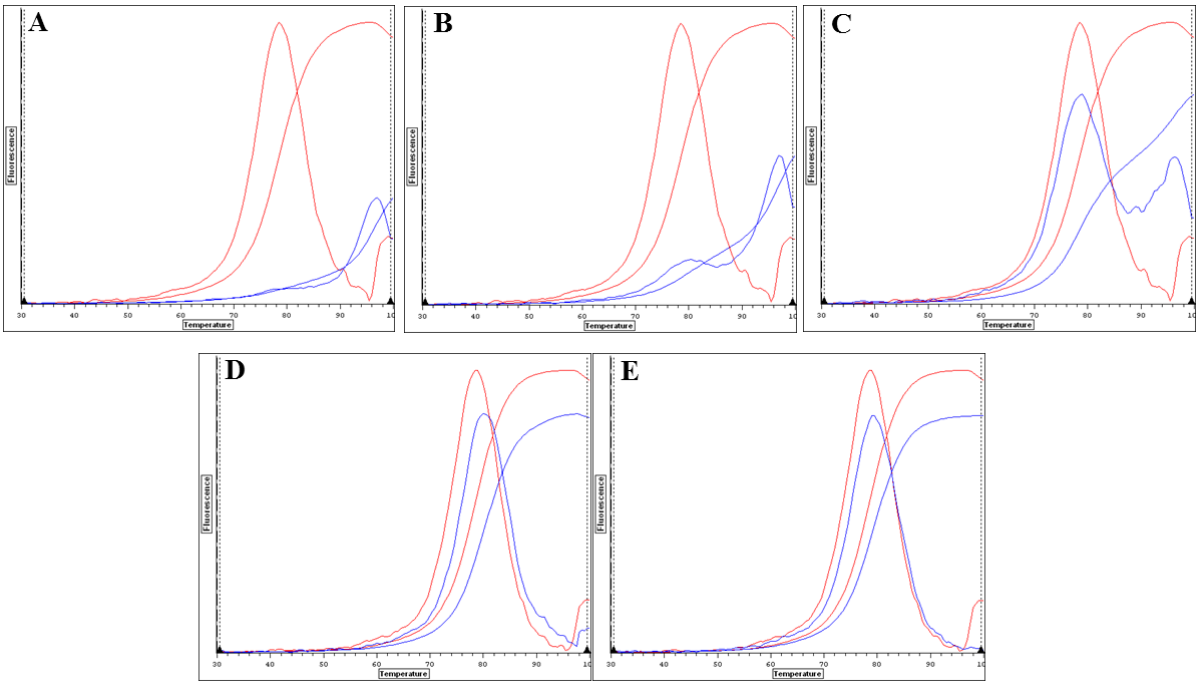


Figure A3.11: FRET melting curves for NF-κB-1 sequence and SJG-136 after a 24 hours incubation. Red curve: NF-κB-1 sequence; blue curve: SJG-136/NF-κB-1 adduct. **A**, SJG-136 = 100 µM; **B**, SJG-136 = 50 µM; **C**, SJG-136 = 10 µM; **D**, SJG-136 = 1 µM; **E**, SJG-136 = 0.1 µM. Refers to Chapter 3, p 129 in thesis.

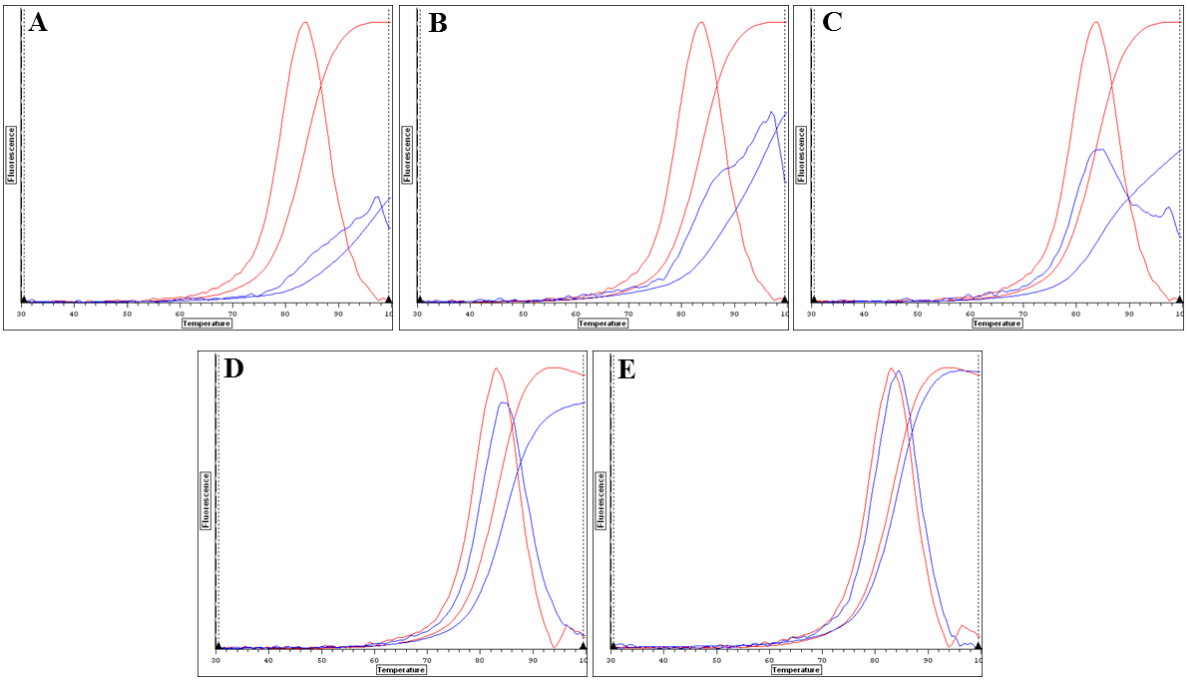


Figure A3.12: FRET melting curves for NF-κB-2 sequence and SJG-136 after a 24 hours incubation. Red curve: NF-κB-2 sequence; blue curve: SJG-136/NF-κB-2 adduct. **A**, SJG-136 = 100 µM; **B**, SJG-136 = 50 µM; **C**, SJG-136 = 10 µM; **D**, SJG-136 = 1 µM; **E**, SJG-136 = 0.1 µM. Refers to Chapter 3, p 129 in thesis.

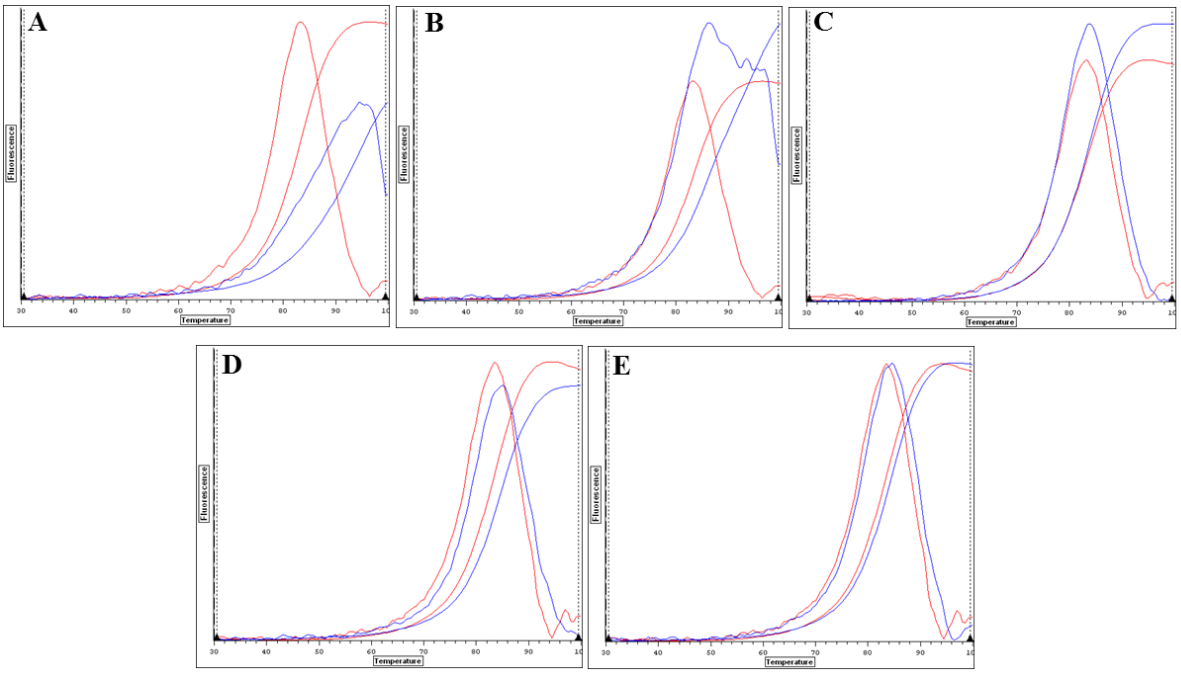


Figure A3.13: FRET melting curves for EGR-1 sequence and SJG-136 after a 24 hours incubation. Red curve: EGR-1 sequence; blue curve: SJG-136/EGR-1 adduct. **A**, SJG-136 = 100 µM; **B**, SJG-136 = 50 µM; **C**, SJG-136 = 10 µM; **D**, SJG-136 = 1 µM; **E**, SJG-136 = 0.1 µM. Refers to Chapter 3, p 129 in thesis.

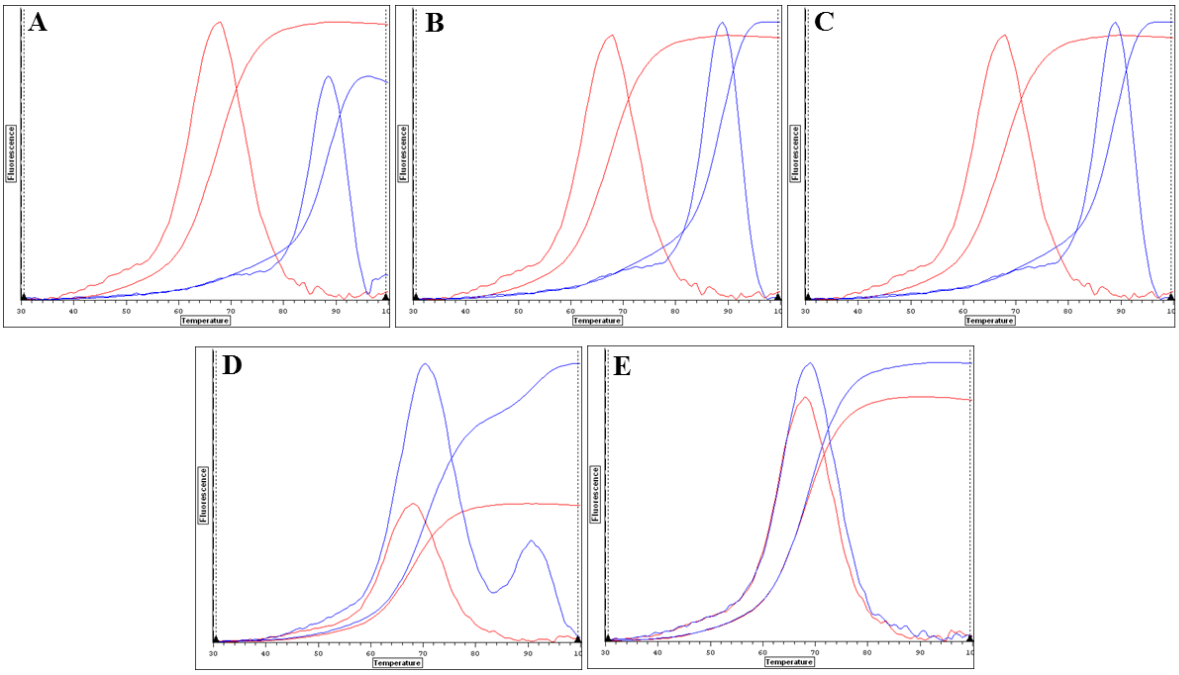


Figure A3.14: FRET melting curves for AP-1 sequence and SJG-136 after a 24 hours incubation. Red curve: AP-1 sequence; blue curve: SJG-136/AP-1 adduct. **A**, SJG-136 = 100 µM; **B**, SJG-136 = 50 µM; **C**, SJG-136 = 10 µM; **D**, SJG-136 = 1 µM; **E**, SJG-136 = 0.1 µM. Refers to Chapter 3, p 129 in thesis.

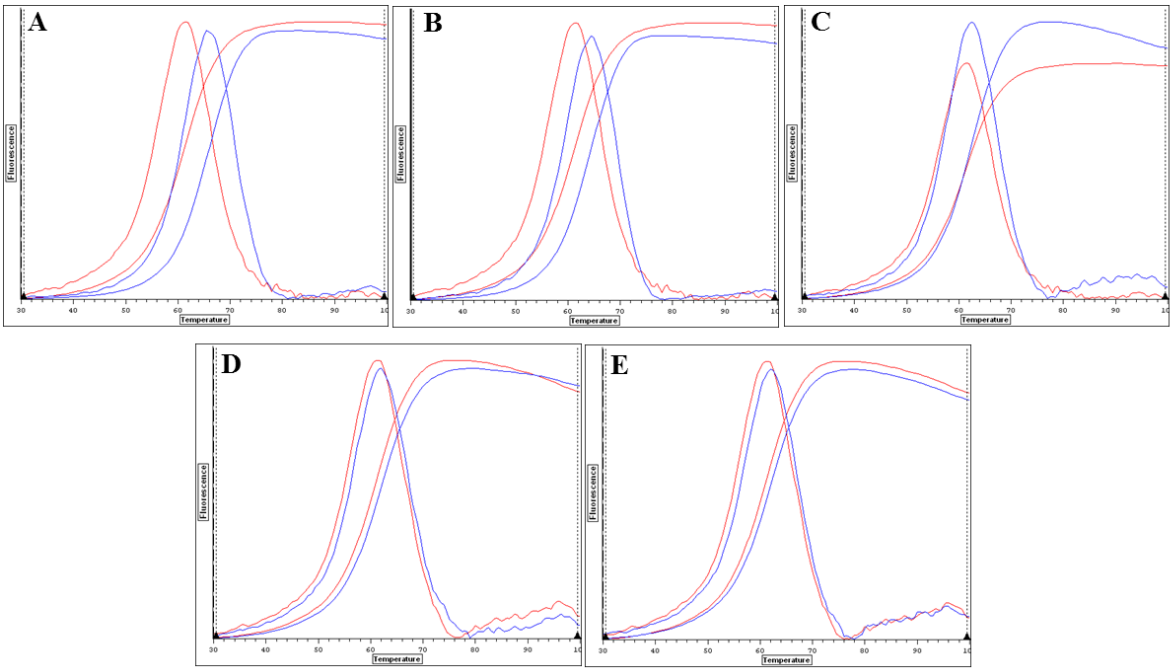


Figure A3.15: FRET melting curves for STAT3 sequence and SJG-136 after a 24 hours incubation. Red curve: STAT3 sequence; blue curve: SJG-136/STAT3 adduct. **A**, SJG-136 = 100 µM; **B**, SJG-136 = 50 µM; **C**, SJG-136 = 10 µM; **D**, SJG-136 = 1 µM; **E**, SJG-136 = 0.1 µM. Refers to Chapter 3, p 129 in thesis.

Bibliography

1. Weinberg, R. A., *The biology of cancer*. Garland Science: New York, 2007.
2. Ferlay, J.; Shin, H. R.; Bray, F.; Forman, D.; Mathers, C.; Parkin, D. M., Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *International journal of cancer. Journal international du cancer* **2010,** *127* (12), 2893-917.
3. Jemal, A.; Bray, F.; Center, M. M.; Ferlay, J.; Ward, E.; Forman, D., Global cancer statistics. *CA: a cancer journal for clinicians* **2011,** *61* (2), 69-90.
4. Jemal, A.; Center, M. M.; DeSantis, C.; Ward, E. M., Global patterns of cancer incidence and mortality rates and trends. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **2010,** *19* (8), 1893-907.
5. Greenwald, P.; Dunn, B. K., Landmarks in the history of cancer epidemiology. *Cancer research* **2009,** *69* (6), 2151-62.
6. National-Academy-of-Sciences Report on Diet and Health. *Nutr Rev* **1989,** *47* (5), 142-149.
7. McDonald, J. C.; McDonald, A. D., The epidemiology of mesothelioma in historical context. *The European respiratory journal* **1996,** *9* (9), 1932-42.
8. Creech, J. L., Jr.; Johnson, M. N., Angiosarcoma of liver in the manufacture of polyvinyl chloride. *Journal of occupational medicine. : official publication of the Industrial Medical Association* **1974,** *16* (3), 150-1.
9. Olfert, S. M.; Felknor, S. A.; Delclos, G. L., An updated review of the literature: risk factors for bladder cancer with focus on occupational exposures. *Southern medical journal* **2006,** *99* (11), 1256-63.
10. (a) Brenner, H.; Bode, G.; Boeing, H., Helicobacter pylori infection among offspring of patients with stomach cancer. *Gastroenterology* **2000,** *118* (1), 31-5; (b) Parkin, D. M., The global health burden of infection-associated cancers in the year 2002. *International journal of cancer. Journal international du cancer* **2006,** *118* (12), 3030-44.
11. (a) zur Hausen, H., Papillomaviruses--to vaccination and beyond. *Biochemistry. Biokhimiia* **2008,** *73* (5), 498-503; (b) Steben, M.; Duarte-Franco, E., Human papillomavirus infection: epidemiology and pathophysiology. *Gynecologic oncology* **2007,** *107* (2 Suppl 1), S2-5.
12. Henderson, B. E.; Ross, R. K.; Pike, M. C.; Casagrande, J. T., Endogenous hormones as a major factor in human cancer. *Cancer research* **1982,** *42* (8), 3232-9.
13. Kantor, R. R.; Mattes, M. J.; Lloyd, K. O.; Old, L. J.; Albino, A. P., Biochemical analysis of two cell surface glycoprotein complexes, very common antigen 1 and very common antigen 2. Relationship to very late activation T cell antigens. *The Journal of biological chemistry* **1987,** *262* (31), 15158-65.
14. Croce, C. M., Oncogenes and cancer. *The New England journal of medicine* **2008,** *358* (5), 502-11.
15. Hanahan, D.; Weinberg, R. A., The hallmarks of cancer. *Cell* **2000,** *100* (1), 57-70.
16. Adams, J. M.; Cory, S., The Bcl-2 apoptotic switch in cancer development and therapy. *Oncogene* **2007,** *26* (9), 1324-37.
17. Colombel, M.; Symmans, F.; Gil, S.; O'Toole, K. M.; Chopin, D.; Benson, M.; Olsson, C. A.; Korsmeyer, S.; Buttyan, R., Detection of the apoptosis-suppressing oncoprotein bc1-2 in hormone-refractory human prostate cancers. *The American journal of pathology* **1993,** *143* (2), 390-400.
18. Ramsay, J. A.; From, L.; Kahn, H. J., bcl-2 protein expression in melanocytic neoplasms of the skin. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* **1995,** *8* (2), 150-4.

19. Kitagawa, Y.; Wong, F.; Lo, P.; Elliott, M.; Verburgt, L. M.; Hogg, J. C.; Daya, M., Overexpression of Bcl-2 and mutations in p53 and K-ras in resected human non-small cell lung cancers. *American journal of respiratory cell and molecular biology* **1996,** *15* (1), 45-54.
20. O'Kane, S. L.; Pound, R. J.; Campbell, A.; Chaudhuri, N.; Lind, M. J.; Cawkwell, L., Expression of bcl-2 family members in malignant pleural mesothelioma. *Acta oncologica* **2006,** *45* (4), 449-53.
21. Fuster, J. J.; Sanz-Gonzalez, S. M.; Moll, U. M.; Andres, V., Classic and novel roles of p53: prospects for anticancer therapy. *Trends in molecular medicine* **2007,** *13* (5), 192-9.
22. (a) Nagy, J. A.; Chang, S. H.; Shih, S. C.; Dvorak, A. M.; Dvorak, H. F., Heterogeneity of the tumor vasculature. *Seminars in thrombosis and hemostasis* **2010,** *36* (3), 321-31; (b) Baluk, P.; Hashizume, H.; McDonald, D. M., Cellular abnormalities of blood vessels as targets in cancer. *Current opinion in genetics & development* **2005,** *15* (1), 102-11.
23. (a) Mattes, J. A., A controlled evaluation of a JCAH regulation. *The Psychiatric hospital* **1987,** *18* (3), 131-4; (b) Folkman, J., Role of angiogenesis in tumor growth and metastasis. *Seminars in oncology* **2002,** *29* (6 Suppl 16), 15-8.
24. Kohn, K. W.; Hartley, J. A.; Mattes, W. B., Mechanisms of DNA sequence selective alkylation of guanine-N7 positions by nitrogen mustards. *Nucleic acids research* **1987,** *15* (24), 10531-49.
25. Osborn, M. J.; Huennekens, F. M., Enzymatic reduction of dihydrofolic acid. *The Journal of biological chemistry* **1958,** *233* (4), 969-74.
26. Johnson, I. S.; Armstrong, J. G.; Gorman, M.; Burnett, J. P., Jr., The Vinca Alkaloids: A New Class of Oncolytic Agents. *Cancer research* **1963,** *23*, 1390-427.
27. Rowinsky, E. K., Current developments in antitumor antibiotics, epipodophyllotoxins, and vinca alkaloids. *Current opinion in oncology* **1991,** *3* (6), 1060-9.
28. Rowinsky, E. K.; Donehower, R. C., Drug-Therapy - Paclitaxel (Taxol). *New Engl J Med* **1995,** *332* (15), 1004-1014.
29. Lord, C. J.; Ashworth, A., Biology-driven cancer drug development: back to the future. *BMC biology* **2010,** *8*, 38.
30. Cole, M. P.; Jones, C. T.; Todd, I. D., A new anti-oestrogenic agent in late breast cancer. An early clinical appraisal of ICI46474. *British journal of cancer* **1971,** *25* (2), 270-5.
31. Sawyers, C., Targeted cancer therapy. *Nature* **2004,** *432* (7015), 294-297.
32. Scott, A. M.; Wolchok, J. D.; Old, L. J., Antibody therapy of cancer. *Nature reviews. Cancer* **2012,** *12* (4), 278-87.
33. Mekapati, S. B.; Denny, W. A.; Kurup, A.; Hansch, C., QSAR of anticancer compounds. Bis(11-oxo-11H-indeno[1,2-b]quinoline-6-carboxamides), bis(phenazine-1-carboxamides), and bis(naphthalimides). *Bioorganic & medicinal chemistry* **2001,** *9* (11), 2757-62.
34. Garnock-Jones, K. P.; Keating, G. M.; Scott, L. J., Trastuzumab: A review of its use as adjuvant treatment in human epidermal growth factor receptor 2 (HER2)-positive early breast cancer. *Drugs* **2010,** *70* (2), 215-39.
35. Yarden, Y.; Sliwkowski, M. X., Untangling the ErbB signalling network. *Nature reviews. Molecular cell biology* **2001,** *2* (2), 127-37.
36. (a) Vaquerizas, J. M.; Kummerfeld, S. K.; Teichmann, S. A.; Luscombe, N. M., A census of human transcription factors: function, expression and evolution. *Nature reviews. Genetics* **2009,** *10* (4), 252-63; (b) Latchman, D. S., Transcription factors: An overview. *The International Journal of Biochemistry & Cell Biology* **1997,** *29* (12), 1305-1312.
37. Boyadjiev, S. A.; Jabs, E. W., Online Mendelian Inheritance in Man (OMIM) as a knowledgebase for human developmental disorders. *Clinical genetics* **2000,** *57* (4), 253-66.
38. Siebenlist, U.; Franzoso, G.; Brown, K., Structure, regulation and function of NF-kappa B. *Annual review of cell biology* **1994,** *10*, 405-55.

39. (a) Moynagh, P. N., The NF-kappaB pathway. *Journal of cell science* **2005,** *118* (Pt 20), 4589-92; (b) Hoffmann, A.; Natoli, G.; Ghosh, G., Transcriptional regulation via the NF-kappaB signaling module. *Oncogene* **2006,** *25* (51), 6706-16.
40. Gilmore, T. D., Introduction to NF-kappaB: players, pathways, perspectives. *Oncogene* **2006,** *25* (51), 6680-4.
41. Chen, F. E.; Huang, D. B.; Chen, Y. Q.; Ghosh, G., Crystal structure of p50/p65 heterodimer of transcription factor NF-kappaB bound to DNA. *Nature* **1998,** *391* (6665), 410-3.
42. Perkins, N. D., Integrating cell-signalling pathways with NF-kappaB and IKK function. *Nature reviews. Molecular cell biology* **2007,** *8* (1), 49-62.
43. Bonizzi, G.; Karin, M., The two NF-kappaB activation pathways and their role in innate and adaptive immunity. *Trends in immunology* **2004,** *25* (6), 280-8.
44. Ghosh, S.; Karin, M., Missing pieces in the NF-kappaB puzzle. *Cell* **2002,** *109 Suppl*, S81-96.
45. (a) Basseres, D. S.; Baldwin, A. S., Nuclear factor-kappaB and inhibitor of kappaB kinase pathways in oncogenic initiation and progression. *Oncogene* **2006,** *25* (51), 6817-30; (b) Jost, P. J.; Ruland, J., Aberrant NF-kappaB signaling in lymphoma: mechanisms, consequences, and therapeutic implications. *Blood* **2007,** *109* (7), 2700-7.
46. Dohner, H.; Stilgenbauer, S.; Benner, A.; Leupolt, E.; Krober, A.; Bullinger, L.; Dohner, K.; Bentz, M.; Lichter, P., Genomic aberrations and survival in chronic lymphocytic leukemia. *The New England journal of medicine* **2000,** *343* (26), 1910-6.
47. (a) Thiel, G.; Cibelli, G., Regulation of life and death by the zinc finger transcription factor Egr-1. *Journal of cellular physiology* **2002,** *193* (3), 287-92; (b) Hallahan, D. E.; Dunphy, E.; Virudachalam, S.; Sukhatme, V. P.; Kufe, D. W.; Weichselbaum, R. R., C-jun and Egr-1 participate in DNA synthesis and cell survival in response to ionizing radiation exposure. *The Journal of biological chemistry* **1995,** *270* (51), 30303-9.
48. Gashler, A.; Sukhatme, V. P., Early growth response protein 1 (Egr-1): prototype of a zinc-finger family of transcription factors. *Progress in nucleic acid research and molecular biology* **1995,** *50*, 191-224.
49. (a) Pavletich, N. P.; Pabo, C. O., Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. *Science* **1991,** *252* (5007), 809-17; (b) Christy, B.; Nathans, D., DNA binding site of the growth factor-inducible protein Zif268. *Proceedings of the National Academy of Sciences of the United States of America* **1989,** *86* (22), 8737-41.
50. (a) Cao, X. M.; Koski, R. A.; Gashler, A.; McKiernan, M.; Morris, C. F.; Gaffney, R.; Hay, R. V.; Sukhatme, V. P., Identification and characterization of the Egr-1 gene product, a DNA-binding zinc finger protein induced by differentiation and growth signals. *Molecular and cellular biology* **1990,** *10* (5), 1931-9; (b) Yan, S. F.; Lu, J.; Zou, Y. S.; Soh-Won, J.; Cohen, D. M.; Buttrick, P. M.; Cooper, D. R.; Steinberg, S. F.; Mackman, N.; Pinsky, D. J.; Stern, D. M., Hypoxia-associated induction of early growth response-1 gene expression. *The Journal of biological chemistry* **1999,** *274* (21), 15030-40.
51. Baron, V.; Adamson, E. D.; Calogero, A.; Ragona, G.; Mercola, D., The transcription factor Egr1 is a direct regulator of multiple tumor suppressors including TGFbeta1, PTEN, p53, and fibronectin. *Cancer gene therapy* **2006,** *13* (2), 115-24.
52. (a) Levin, W. J.; Press, M. F.; Gaynor, R. B.; Sukhatme, V. P.; Boone, T. C.; Reissmann, P. T.; Figlin, R. A.; Holmes, E. C.; Souza, L. M.; Slamon, D. J., Expression patterns of immediate early transcription factors in human non-small cell lung cancer. The Lung Cancer Study Group. *Oncogene* **1995,** *11* (7), 1261-9; (b) Joslin, J. M.; Fernald, A. A.; Tennant, T. R.; Davis, E. M.; Kogan, S. C.; Anastasi, J.; Crispino, J. D.; Le Beau, M. M., Haploinsufficiency of EGR1, a candidate gene in the del(5q), leads to the development of myeloid disorders. *Blood* **2007,** *110* (2), 719-26; (c) Huang, R. P.; Fan, Y.; de Belle, I.; Niemeyer, C.; Gottardis, M. M.; Mercola, D.; Adamson, E. D., Decreased Egr-1 expression in human, mouse and rat mammary cells and tissues correlates with tumor formation. *International journal of cancer. Journal international*

du cancer **1997,** *72* (1), 102-9; (d) Calogero, A.; Arcella, A.; De Gregorio, G.; Porcellini, A.; Mercola, D.; Liu, C.; Lombari, V.; Zani, M.; Giannini, G.; Gagliardi, F. M.; Caruso, R.; Gulino, A.; Frati, L.; Ragona, G., The early growth response gene EGR-1 behaves as a suppressor gene that is down-regulated independent of ARF/Mdm2 but not p53 alterations in fresh human gliomas. *Clinical cancer research : an official journal of the American Association for Cancer Research* **2001,** *7* (9), 2788-96.
53. (a) Angel, P.; Karin, M., The role of Jun, Fos and the AP-1 complex in cell-proliferation and transformation. *Biochimica et biophysica acta* **1991,** *1072* (2-3), 129-57; (b) Vogt, P. K.; Bos, T. J., jun: oncogene and transcription factor. *Advances in cancer research* **1990,** *55*, 1-35.
54. Zhang, W.; Hart, J.; McLeod, H. L.; Wang, H. L., Differential expression of the AP-1 transcription factor family members in human colorectal epithelial and neuroendocrine neoplasms. *American journal of clinical pathology* **2005,** *124* (1), 11-9.
55. (a) Halazonetis, T. D.; Georgopoulos, K.; Greenberg, M. E.; Leder, P., c-Jun dimerizes with itself and with c-Fos, forming complexes of different DNA binding affinities. *Cell* **1988,** *55* (5), 917-24; (b) Smeal, T.; Angel, P.; Meek, J.; Karin, M., Different requirements for formation of Jun: Jun and Jun: Fos complexes. *Genes & development* **1989,** *3* (12B), 2091-100.
56. Angel, P.; Imagawa, M.; Chiu, R.; Stein, B.; Imbra, R. J.; Rahmsdorf, H. J.; Jonat, C.; Herrlich, P.; Karin, M., Phorbol ester-inducible genes contain a common cis element recognized by a TPA-modulated trans-acting factor. *Cell* **1987,** *49* (6), 729-39.
57. Chinenov, Y.; Kerppola, T. K., Close encounters of many kinds: Fos-Jun interactions that mediate transcription regulatory specificity. *Oncogene* **2001,** *20* (19), 2438-52.
58. Hess, J.; Angel, P.; Schorpp-Kistner, M., AP-1 subunits: quarrel and harmony among siblings. *Journal of cell science* **2004,** *117* (Pt 25), 5965-73.
59. (a) Shaulian, E.; Karin, M., AP-1 in cell proliferation and survival. *Oncogene* **2001,** *20* (19), 2390-400; (b) Shaulian, E.; Karin, M., AP-1 as a regulator of cell life and death. *Nat Cell Biol* **2002,** *4* (5), E131-E136; (c) Wagner, E. F.; Eferl, R., Fos/AP-1 proteins in bone and the immune system. *Immunological reviews* **2005,** *208*, 126-40.
60. (a) Mathas, S.; Hinz, M.; Anagnostopoulos, I.; Krappmann, D.; Lietz, A.; Jundt, F.; Bommert, K.; Mechta-Grigoriou, F.; Stein, H.; Dorken, B.; Scheidereit, C., Aberrantly expressed c-Jun and JunB are a hallmark of Hodgkin lymphoma cells, stimulate proliferation and synergize with NF-kappa B. *The EMBO journal* **2002,** *21* (15), 4104-13; (b) Milde-Langosch, K., The Fos family of transcription factors and their role in tumourigenesis. *European journal of cancer* **2005,** *41* (16), 2449-61.
61. (a) Heinrich, P. C.; Behrmann, I.; Muller-Newen, G.; Schaper, F.; Graeve, L., Interleukin-6-type cytokine signalling through the gp130/Jak/STAT pathway. *The Biochemical journal* **1998,** *334 (Pt 2)*, 297-314; (b) Leeman, R. J.; Lui, V. W.; Grandis, J. R., STAT3 as a therapeutic target in head and neck cancer. *Expert opinion on biological therapy* **2006,** *6* (3), 231-41; (c) Quesnelle, K. M.; Boehm, A. L.; Grandis, J. R., STAT-mediated EGFR signaling in cancer. *Journal of cellular biochemistry* **2007,** *102* (2), 311-9.
62. Becker, S.; Groner, B.; Muller, C. W., Three-dimensional structure of the Stat3beta homodimer bound to DNA. *Nature* **1998,** *394* (6689), 145-51.
63. (a) Germain, D.; Frank, D. A., Targeting the cytoplasmic and nuclear functions of signal transducers and activators of transcription 3 for cancer therapy. *Clinical cancer research : an official journal of the American Association for Cancer Research* **2007,** *13* (19), 5665-9; (b) Heinrich, P. C.; Behrmann, I.; Haan, S.; Hermanns, H. M.; Muller-Newen, G.; Schaper, F., Principles of interleukin (IL)-6-type cytokine signalling and its regulation. *The Biochemical journal* **2003,** *374* (Pt 1), 1-20.
64. Johnston, P. A.; Grandis, J. R., STAT3 signaling: anticancer strategies and challenges. *Molecular interventions* **2011,** *11* (1), 18-26.
65. (a) Jing, N.; Tweardy, D. J., Targeting Stat3 in cancer therapy. *Anti-cancer drugs* **2005,** *16* (6), 601-7; (b) Niu, G.; Wright, K. L.; Huang, M.; Song, L.; Haura, E.; Turkson, J.; Zhang, S.;

Wang, T.; Sinibaldi, D.; Coppola, D.; Heller, R.; Ellis, L. M.; Karras, J.; Bromberg, J.; Pardoll, D.; Jove, R.; Yu, H., Constitutive Stat3 activity up-regulates VEGF expression and tumor angiogenesis. *Oncogene* **2002,** *21* (13), 2000-8.
66. German, C. L.; Sauer, B. M.; Howe, C. L., The STAT3 beacon: IL-6 recurrently activates STAT 3 from endosomal structures. *Experimental Cell Research* **2011,** *317* (14), 1955-1969.
67. Yakovchuk, P.; Protozanova, E.; Frank-Kamenetskii, M. D., Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic acids research* **2006,** *34* (2), 564-574.
68. Wing, R.; Drew, H.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E., Crystal structure analysis of a complete turn of B-DNA. *Nature* **1980,** *287* (5784), 755-8.
69. Hurley, L. H., DNA and its associated processes as targets for cancer therapy. *Nature reviews. Cancer* **2002,** *2* (3), 188-200.
70. Palchaudhuri, R.; Hergenrother, P. J., DNA as a target for anticancer compounds: methods to determine the mode of binding and the mechanism of action. *Curr Opin Biotech* **2007,** *18* (6), 497-503.
71. Hartley, J. A.; Hochhauser, D., Small molecule drugs - optimizing DNA damaging agent-based therapeutics. *Current opinion in pharmacology* **2012,** *12* (4), 398-402.
72. Faulds, D.; Balfour, J. A.; Chrisp, P.; Langtry, H. D., Mitoxantrone - a Review of Its Pharmacodynamic and Pharmacokinetic Properties, and Therapeutic Potential in the Chemotherapy of Cancer. *Drugs* **1991,** *41* (3), 400-449.
73. Young, R. C.; Ozols, R. F.; Myers, C. E., The anthracycline antineoplastic drugs. *The New England journal of medicine* **1981,** *305* (3), 139-53.
74. (a) Reddy, B. S.; Sondhi, S. M.; Lown, J. W., Synthetic DNA minor groove-binding drugs. *Pharmacology & therapeutics* **1999,** *84* (1), 1-111; (b) Ren, J.; Chaires, J. B., Sequence and structural selectivity of nucleic acid binding ligands. *Biochemistry* **1999,** *38* (49), 16067-75.
75. Black, D. J.; Livingston, R. B., Antineoplastic drugs in 1990. A review (Part II). *Drugs* **1990,** *39* (5), 652-73.
76. Denny, W. A., DNA minor groove alkylating agents. *Current medicinal chemistry* **2001,** *8* (5), 533-44.
77. Saffhill, R.; Margison, G. P.; O'Connor, P. J., Mechanisms of carcinogenesis induced by alkylating agents. *Biochimica et biophysica acta* **1985,** *823* (2), 111-45.
78. McHugh, P. J.; Spanswick, V. J.; Hartley, J. A., Repair of DNA interstrand crosslinks: molecular mechanisms and clinical relevance. *The Lancet. Oncology* **2001,** *2* (8), 483-90.
79. Goodman, L. S.; Wintrobe, M. M.; Dameshek, W.; Goodman, M. J.; Gilman, A.; McLennan, M. T., Landmark article Sept. 21, 1946: Nitrogen mustard therapy. Use of methyl-bis(beta-chloroethyl)amine hydrochloride and tris(beta-chloroethyl)amine hydrochloride for Hodgkin's disease, lymphosarcoma, leukemia and certain allied and miscellaneous disorders. By Louis S. Goodman, Maxwell M. Wintrobe, William Dameshek, Morton J. Goodman, Alfred Gilman and Margaret T. McLennan. *Jama* **1984,** *251* (17), 2255-61.
80. Wong, E.; Giandomenico, C. M., Current status of platinum-based antitumor drugs. *Chemical reviews* **1999,** *99* (9), 2451-66.
81. Turner, P. R.; Denny, W. A., The genome as a drug target: sequence specific minor groove binding ligands. *Current drug targets* **2000,** *1* (1), 1-14.
82. Wemmer, D. E.; Dervan, P. B., Targeting the minor groove of DNA. *Current opinion in structural biology* **1997,** *7* (3), 355-61.
83. Dervan, P. B., Molecular recognition of DNA by small molecules. *Bioorganic & medicinal chemistry* **2001,** *9* (9), 2215-35.
84. Wellenzohn, B.; Flader, W.; Winger, R. H.; Hallbrucker, A.; Mayer, E.; Liedl, K. R., Significance of ligand tails for interaction with the minor groove of B-DNA. *Biophysical journal* **2001,** *81* (3), 1588-99.

85. Baraldi, P. G.; Bovero, A.; Fruttarolo, F.; Preti, D.; Tabrizi, M. A.; Pavani, M. G.; Romagnoli, R., DNA minor groove binders as potential antitumor and antimicrobial agents. *Medicinal research reviews* **2004,** *24* (4), 475-528.
86. Van Dyke, M. W.; Hertzberg, R. P.; Dervan, P. B., Map of distamycin, netropsin, and actinomycin binding sites on heterogeneous DNA: DNA cleavage-inhibition patterns with methidiumpropyl-EDTA.Fe(II). *Proceedings of the National Academy of Sciences of the United States of America* **1982,** *79* (18), 5470-4.
87. Abu-Daya, A.; Brown, P. M.; Fox, K. R., DNA sequence preferences of several AT-selective minor groove binding ligands. *Nucleic acids research* **1995,** *23* (17), 3385-92.
88. (a) Neidle, S., DNA minor-groove recognition by small molecules. *Natural product reports* **2001,** *18* (3), 291-309; (b) Smellie, M.; Bose, D. S.; Thompson, A. S.; Jenkins, T. C.; Hartley, J. A.; Thurston, D. E., Sequence-Selective Recognition of Duplex DNA Through Covalent Interstrand Cross-Linking: Kinetic and Molecular Modeling Studies With Pyrrolobenzodiazepine Dimers. *Biochemistry* **2003,** *42* (27), 8232-8239.
89. (a) Antonow, D.; Thurston, D. E., Synthesis of DNA-interactive pyrrolo[2,1-c][1,4]benzodiazepines (PBDs). *Chemical reviews* **2011,** *111* (4), 2815-64; (b) Wells, G.; Martin, C. R.; Howard, P. W.; Sands, Z. A.; Laughton, C. A.; Tiberghien, A.; Woo, C. K.; Masterson, L. A.; Stephenson, M. J.; Hartley, J. A.; Jenkins, T. C.; Shnyder, S. D.; Loadman, P. M.; Waring, M. J.; Thurston, D. E., Design, synthesis, and biophysical and biological evaluation of a series of pyrrolobenzodiazepine-poly(N-methylpyrrole) conjugates. *J Med Chem* **2006,** *49* (18), 5442-61.
90. Rahman, K. M.; Jackson, P. J.; James, C. H.; Basu, B. P.; Hartley, J. A.; de la Fuente, M.; Schatzlein, A.; Robson, M.; Pedley, R. B.; Pepper, C.; Fox, K. R.; Howard, P. W.; Thurston, D. E., GC-targeted C8-linked pyrrolobenzodiazepine-biaryl conjugates with femtomolar in vitro cytotoxicity and in vivo antitumor activity in mouse models. *J Med Chem* **2013,** *56* (7), 2911-35.
91. Rahman, K. M.; James, C. H.; Thurston, D. E., Observation of the reversibility of a covalent pyrrolobenzodiazepine (PBD) DNA adduct by HPLC/MS and CD spectroscopy. *Organic & biomolecular chemistry* **2011,** *9* (5), 1632-41.
92. Leimgruber, W.; Stefanovic, V.; Schenker, F.; Karr, A.; Berger, J., Isolation and characterization of anthramycin, a new antitumor antibiotic. *J Am Chem Soc* **1965,** *87* (24), 5791-3.
93. (a) Farmer, J. D.; Rudnicki, S. M.; Suggs, J. W., Synthesis and DNA Crosslinking Ability of a Dimeric Anthramycin Analog. *Tetrahedron Letters* **1988,** *29* (40), 5105-5108; (b) Farmer, J. D., Jr.; Gustafson, G. R.; Conti, A.; Zimmt, M. B.; Suggs, J. W., DNA binding properties of a new class of linked anthramycin analogs. *Nucleic acids research* **1991,** *19* (4), 899-903.
94. Gregson, S. J.; Howard, P. W.; Hartley, J. A.; Brooks, N. A.; Adams, L. J.; Jenkins, T. C.; Kelland, L. R.; Thurston, D. E., Design, synthesis, and evaluation of a novel pyrrolobenzodiazepine DNA-interactive agent with highly efficient cross-linking ability and potent cytotoxicity. *J Med Chem* **2001,** *44* (5), 737-48.
95. (a) Gregson, S. J.; Howard, P. W.; Thurston, D. E., Synthesis of the first examples of A-C8/C-C2 amide-Linked pyrrolo[2,1-c][1,4]benzodiazepine dimers. *Bioorganic & Medicinal Chemistry Letters* **2003,** *13* (14), 2277-2280; (b) Reddy, B. S.; Damayanthi, Y.; Lown, J. W., Design, synthesis and in vitro cytotoxicity studies of novel pyrrolo[2,1-c][1,4]benzodiazepine (PBD)--polymade conjugates and 2,2'-PBD dimers. *Anticancer Drug Des* **2000,** *15* (3), 225-38.
96. Rahman, K. M.; Thompson, A. S.; James, C. H.; Narayanaswamy, M.; Thurston, D. E., The pyrrolobenzodiazepine dimer SJG-136 forms sequence-dependent intrastrand DNA cross-links and monoalkylated adducts in addition to interstrand cross-links. *Journal of the American Chemical Society* **2009,** *131* (38), 13756-66.
97. Gregson, S. J.; Howard, P. W.; Gullick, D. R.; Hamaguchi, A.; Corcoran, K. E.; Brooks, N. A.; Hartley, J. A.; Jenkins, T. C.; Patel, S.; Guille, M. J.; Thurston, D. E., Linker length modulates DNA cross-linking reactivity and cytotoxic potency of C8/C8' ether-linked C2-exo-unsaturated pyrrolo[2,1-c][1,4]benzodiazepine (PBD) dimers. *J Med Chem* **2004,** *47* (5), 1161-74.

98. (a) Janjigian, Y. Y.; Lee, W.; Kris, M. G.; Miller, V. A.; Krug, L. M.; Azzoli, C. G.; Senturk, E.; Calcutt, M. W.; Rizvi, N. A., A Phase I Trial of SJG-136 (NSC#694501) in Advanced Solid Tumors. *Cancer Chemotherapy and Pharmacology* **2010,** *65* (5), 833-838; (b) Hochhauser, D.; Meyer, T.; Spanswick, V. J.; Wu, J.; Clingen, P. H.; Loadman, P.; Cobb, M.; Gumbrell, L.; Begent, R. H.; Hartley, J. A.; Jodrell, D., Phase I study of sequence-selective minor groove DNA binding agent SJG-136 in patients with advanced solid tumors. *Clinical cancer research : an official journal of the American Association for Cancer Research* **2009,** *15* (6), 2140-7.
99. Kung Sutherland, M. S.; Walter, R. B.; Jeffrey, S. C.; Burke, P. J.; Yu, C.; Kostner, H.; Stone, I.; Ryan, M. C.; Sussman, D.; Lyon, R. P.; Zeng, W.; Harrington, K. H.; Klussman, K.; Westendorf, L.; Meyer, D.; Bernstein, I. D.; Senter, P. D.; Benjamin, D. R.; Drachman, J. G.; McEarchern, J. A., SGN-CD33A: a novel CD33-targeting antibody-drug conjugate using a pyrrolobenzodiazepine dimer is active in models of drug-resistant AML. *Blood* **2013,** *122* (8), 1455-63.
100. Puvvada, M. S.; Hartley, J. A.; Jenkins, T. C.; Thurston, D. E., A Quantitative Assay to Measure the Relative DNA-Binding Affinity of Pyrrolo[2,1-*c*][1,4]benzodiazepine (PBD) Antitumor Antibiotics Based on the Inhibition of Restriction-Endonuclease BamHI. *Nucleic acids research* **1993,** *21* (16), 3671-3675.
101. Puvvada, M. S.; Forrow, S. A.; Hartley, J. A.; Stephenson, P.; Gibson, I.; Jenkins, T. C.; Thurston, D. E., Inhibition of Bacteriophage T7 RNA Polymerase *In Vitro* Transcription by DNA-Binding Pyrrolo[2,1-*c*][1,4]benzodiazepines. *Biochemistry* **1997,** *36* (9), 2478-2484.
102. Kotecha, M.; Kluza, J.; Wells, G.; O'Hare, C. C.; Forni, C.; Mantovani, R.; Howard, P. W.; Morris, P.; Thurston, D. E.; Hartley, J. A.; Hochhauser, D., Inhibition of DNA binding of the NF-Y transcription factor by the pyrrolobenzodiazepine-polyamide conjugate GWL-78. *Mol Cancer Ther* **2008,** *7* (5), 1319-28.
103. (a) Kohn, K. W.; Spears, C. L., Reaction of anthramycin with deoxyribonucleic acid. *J Mol Biol* **1970,** *51* (3), 551-72; (b) Kohn, K. W.; Glaubiger, D.; Spears, C. L., The reaction of anthramycin with DNA. II. Studies of kinetics and mechanism. *Biochim Biophys Acta* **1974,** *361* (3), 288-302.
104. Hurley, L. H.; Gairola, C.; Zmijewski, M., Pyrrolo(1,4)benzodiazepine antitumor antibiotics. In vitro interaction of anthramycin, sibiromycin and tomaymycin with DNA using specifically radiolabelled molecules. *Biochim Biophys Acta* **1977,** *475* (3), 521-35.
105. Glaubiger, D.; Kohn, K. W.; Charney, E., The reaction of anthramycin with DNA. 3. Properties of the complex. *Biochim Biophys Acta* **1974,** *361* (3), 303-11.
106. (a) Antonow, D.; Jenkins, T. C.; Howard, P. W.; Thurston, D. E., Synthesis of a novel C2-aryl pyrrolo[2,1-c][1,4]benzodiazepine-5,11-dione library: Effect of C2-aryl substitution on cytotoxicity and non-covalent DNA binding. *Bioorganic & medicinal chemistry* **2007,** *15* (8), 3041-3053; (b) Kamal, A.; Ramesh, G.; Laxman, N.; Ramulu, P.; Srinivas, O.; Neelima, K.; Kondapi, A. K.; Sreenu, V. B.; Nagarajaram, H. A., Design, synthesis, and evaluation of new noncross-linking pyrrolobenzodiazepine dimers with efficient DNA binding ability and potent antitumor activity. *J Med Chem* **2002,** *45* (21), 4679-88.
107. Malhotra, R. K.; Ostrander, J. M.; Hurley, L. H.; McInnes, A. G.; Smith, D. G.; Walter, J. A.; Wright, J. L., Chemical conversion of anthramycin 11-methyl ether to didehydroanhydroanthramycin and its utilization in studies of the biosynthesis and mechanism of action of anthramycin. *J Nat Prod* **1981,** *44* (1), 38-44.
108. Reddy, B. S.; Sharma, S. K.; Lown, J. W., Recent developments in sequence selective minor groove DNA effectors. *Current medicinal chemistry* **2001,** *8* (5), 475-508.
109. (a) Graves, D. E.; Stone, M. P.; Krugh, T. R., Structure of the anthramycin-d(ATGCAT)2 adduct from one- and two-dimensional proton NMR experiments in solution. *Biochemistry* **1985,** *24* (26), 7573-81; (b) Krugh, T. R.; Graves, D. E.; Stone, M. P., Two-dimensional NMR studies on the anthramycin-d(ATGCAT)2 adduct. *Biochemistry* **1989,** *28* (26), 9988-94.

110. Kopka, M. L.; Goodsell, D. S.; Baikalov, I.; Grzeskowiak, K.; Cascio, D.; Dickerson, R. E., Crystal structure of a covalent DNA-drug adduct: anthramycin bound to C-C-A-A-C-G-T-T-G-G and a molecular explanation of specificity. *Biochemistry* **1994,** *33* (46), 13593-610.
111. Boyd, F. L.; Stewart, D.; Remers, W. A.; Barkley, M. D.; Hurley, L. H., Characterization of a unique tomaymycin-d(CICGAATTCICG)2 adduct containing two drug molecules per duplex by NMR, fluorescence, and molecular modeling studies. *Biochemistry* **1990,** *29* (9), 2387-403.
112. Hurley, L. H.; Reck, T.; Thurston, D. E.; Langley, D. R.; Holden, K. G.; Hertzberg, R. P.; Hoover, J. R.; Gallagher, G., Jr.; Faucette, L. F.; Mong, S. M.; et al., Pyrrolo[1,4]benzodiazepine antitumor antibiotics: relationship of DNA alkylation and sequence specificity to the biological activity of natural and synthetic compounds. *Chem Res Toxicol* **1988,** *1* (5), 258-68.
113. (a) Jenkins, T. C.; Hurley, L. H.; Neidle, S.; Thurston, D. E., Structure of a covalent DNA minor groove adduct with a pyrrolobenzodiazepine dimer: evidence for sequence-specific interstrand cross-linking. *J Med Chem* **1994,** *37* (26), 4529-37; (b) Hopton, S. R.; Thompson, A. S., Nuclear Magnetic Resonance Solution Structures of Inter- and Intrastrand Adducts of DNA Cross-Linker SJG-136. *Biochemistry* **2011,** *50* (21), 4720-4732.
114. (a) Adams, L. J.; Jenkins, T. C.; Banting, L.; Thurston, D. E., Molecular Modelling of a Sequence-Specific DNA-Binding Agent Based on the Pyrrolo[2,1-*c*][1,4]benzodiazepines. *Pharmacy & Pharmacology Communications* **1999,** *5*, 555-560; (b) Remers, W. A.; Mabilia, M.; Hopfinger, A. J., Conformations of complexes between pyrrolo[1,4]benzodiazepines and DNA segments. *J. Med. Chem.* **1986,** *29* (12), 2492-2503; (c) Jackson, P. J. M.; James, C. H.; Jenkins, T. C.; Rahman, K. M.; Thurston, D. E., Computational Studies Support the Role of the C7-Sibirosamine Sugar of the Pyrrolobenzodiazepine (PBD) Sibiromycin in Transcription Factor Inhibition. *ACS Chemical Biology* **2014,** *9* (10), 2432-2440.
115. Barkley, M. D.; Cheatham, S.; Thurston, D. E.; Hurley, L. H., Pyrrolo[1,4]benzodiazepine antitumor antibiotics: evidence for two forms of tomaymycin bound to DNA. *Biochemistry* **1986,** *25* (10), 3021-3031.
116. Thurston, D. E.; Bose, D. S.; Howard, P. W.; Jenkins, T. C.; Leoni, A.; Baraldi, P. G.; Guiotto, A.; Cacciari, B.; Kelland, L. R.; Foloppe, M. P.; Rault, S., Effect of A-ring modifications on the DNA-binding behavior and cytotoxicity of pyrrolo[2,1-c][1,4]benzodiazepines. *J Med Chem* **1999,** *42* (11), 1951-64.
117. Antonow, D.; Jenkins, T. C.; Howard, P. W.; Thurston, D. E., Synthesis of a novel C2-aryl pyrrolo[2,1-c][1,4]benzodiazepine-5,11-dione library: effect of C2-aryl substitution on cytotoxicity and non-covalent DNA binding. *Bioorganic & medicinal chemistry* **2007,** *15* (8), 3041-53.
118. Wu, J.; Clingen, P. H.; Spanswick, V. J.; Mellinas-Gomez, M.; Meyer, T.; Puzanov, I.; Jodrell, D.; Hochhauser, D.; Hartley, J. A., gamma-H2AX foci formation as a pharmacodynamic marker of DNA damage produced by DNA cross-linking agents: results from 2 phase I clinical trials of SJG-136 (SG2000). *Clinical Cancer Research* **2013,** *19* (3), 721-30.
119. Cargill, C.; Bachmann, E.; Zbinden, G., Effects of daunomycin and anthramycin on electrocardiogram and mitochondrial metabolism of the rat heart. *Journal of the National Cancer Institute* **1974,** *53* (2), 481-6.
120. Lubawy, W. C.; Dallam, R. A.; Hurley, L. H., Protection against anthramycin-induced toxicity in mice by coenzyme Q10. *Journal of the National Cancer Institute* **1980,** *64* (1), 105-9.
121. Arima, K.; Kosaka, M.; Tamura, G.; Imanaka, H.; Sakai, H., Studies on tomaymycin, a new antibiotic. I. Isolation and properties of tomaymycin. *J Antibiot (Tokyo)* **1972,** *25* (8), 437-44.
122. Brazhnikova, M. G.; Konstantinova, N. V.; Mesentsev, A. S., Sibiromycin: isolation and characterization. *J Antibiot (Tokyo)* **1972,** *25* (11), 668-73.
123. Hisamatsu, T.; Uchida, S.; Takeuchi, T.; Ishizuka, M.; Umezawa, H., Antitumor effect of a new antibiotic, neothramycin. *Gann* **1980,** *71* (3), 308-12.

124. (a) Kaneko, T.; Wong, H.; Doyle, T. W., A total synthesis of chicamycin A a new pyrrolo[1,4]benzodiazepine antitumor agent. *J Antibiot (Tokyo)* **1984,** *37* (3), 300-2; (b) Konishi, M.; Hatori, M.; Tomita, K.; Sugawara, M.; Ikeda, C.; Nishiyama, Y.; Imanishi, H.; Miyaki, T.; Kawaguchi, H., Chicamycin, a new antitumor antibiotic. I. Production, isolation and properties. *J Antibiot (Tokyo)* **1984,** *37* (3), 191-9; (c) Konishi, M.; Ohkuma, H.; Naruse, N.; Kawaguchi, H., Chicamycin, a new antitumor antibiotic. II. Structure determination of chicamycins A and B. *The Journal of antibiotics* **1984,** *37* (3), 200-6.
125. Thurston, D. E., Advances in the Study of Pyrrolo[2,1-*c*][1,4]benzodiazepine (PBD) Antitumour Antibiotics. In *Molecular Aspects of Anticancer Drug-DNA Interactions*, Neidle, S.; Waring, M. J., Eds. The Macmillan Press Ltd., London, UK: London, 1993; Vol. 1, pp 54-88.
126. Walton, M. I.; Goddard, P.; Kelland, L. R.; Thurston, D. E.; Harrap, K. R., Preclinical pharmacology and antitumour activity of the novel sequence-selective DNA minor-groove cross-linking agent DSB-120. *Cancer Chemother Pharmacol* **1996,** *38* (5), 431-8.
127. Cipolla, L.; Araujo, A. C.; Airoldi, C.; Bini, D., Pyrrolo[2,1-c][1,4]benzodiazepine as a scaffold for the design and synthesis of anti-tumour drugs. *Anticancer Agents Med Chem* **2009,** *9* (1), 1-31.
128. (a) Borkovec, A. B.; Chang, S. C.; Horwitz, S. B., Chemosterilization of house flies with anthramycin methyl ether. *J Econ Entomol* **1971,** *64* (4), 983-4; (b) Horwitz, S. B.; Chang, S. C.; Grollman, A. P.; Borkovec, A. B., Chemosterilant action of anthramycin: a proposed mechanism. *Science* **1971,** *174* (4005), 159-61.
129. Gerratana, B., Biosynthesis, synthesis, and biological activities of pyrrolobenzodiazepines. *Medicinal research reviews* **2012,** *32* (2), 254-93.
130. Li, W., Fishkin, N. E., Miller, M., Leece, B., Reid, E., Archer, K., Maloney, E., Jones, G., Kovtun, Y., Singh, R., Chari, R. V. J. In *Design, synthesis and evaluation of a novel DNA-interactive agent: A promising new class of cytotoxic molecules for use in antibody-drug conjugates*, 239th ACS National Meeting, San Francisco, California, San Francisco, California, 2013.
131. Polson, A. G.; Calemine-Fenaux, J.; Chan, P.; Chang, W.; Christensen, E.; Clark, S.; de Sauvage, F. J.; Eaton, D.; Elkins, K.; Elliott, J. M.; Frantz, G.; Fuji, R. N.; Gray, A.; Harden, K.; Ingle, G. S.; Kljavin, N. M.; Koeppen, H.; Nelson, C.; Prabhu, S.; Raab, H.; Ross, S.; Stephan, J.-P.; Scales, S. J.; Spencer, S. D.; Vandlen, R.; Wranik, B.; Yu, S.-F.; Zheng, B.; Ebens, A., Antibody-Drug Conjugates for the Treatment of Non–Hodgkin's Lymphoma: Target and Linker-Drug Selection. *Cancer Research* **2009,** *69* (6), 2358-2364.
132. Sackinger, E.; Boser, B. E.; Bromley, J.; Lecun, Y.; Jackel, L. D., Application of the ANNA neural network chip to high-speed character recognition. *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council* **1992,** *3* (3), 498-505.
133. Bose, D. S.; Thompson, A. S.; Smellie, M.; Berardini, M. D.; Hartley, J. A.; Jenkins, T. C.; Neidle, S.; Thurston, D. E., Effect of Linker Length on DNA-Binding Affinity, Cross-Linking Efficiency and Cytotoxicity of C8-Linked Pyrrolobenzodiazepine Dimers. *J Chem Soc Chem Comm* **1992,** (20), 1518-1520.
134. Morris, S. J.; Thurston, D. E.; Nevell, T. G., Evaluation of the electrophilicity of DNA-binding pyrrolo[2,1-c][1,4]benzodiazepines by HPLC. *J Antibiot (Tokyo)* **1990,** *43* (10), 1286-92.
135. Cheung, A.; Struble, E.; He, J.; Yang, C.; Wang, E.; Thurston, D. E.; Liu, P., Direct liquid chromatography determination of the reactive imine SJG-136 (NSC 694501). *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences* **2005,** *822* (1-2), 10-20.
136. Rahman, K. M.; James, C. H.; Thurston, D. E., Effect of base sequence on the DNA cross-linking properties of pyrrolobenzodiazepine (PBD) dimers. *Nucleic acids research* **2011,** *39* (13), 5800-12.
137. Hartley, J. A.; Hochhauser, D., Small molecule drugs - optimizing DNA damaging agent-based therapeutics. *Current opinion in pharmacology* **2012,** *12* (4), 398-402.

138. Kamal, A.; Ramulu, P.; Srinivas, O.; Ramesh, G., Synthesis and DNA-binding affinity of A-C8/C-C2 alkoxyamido-linked pyrrolo[2,1-c][1,4]benzodiazepine dimers. *Bioorg Med Chem Lett* **2003,** *13* (22), 3955-8.
139. Jackson, P. J. M. Development of In Silico Techniques for the Rational Design of Novel Pyrrolobenzodiazepine-Based Transcription Factor Inhibitors. King's College London, London, 2014.
140. Janjigian, Y. Y.; Lee, W.; Kris, M. G.; Miller, V. A.; Krug, L. M.; Azzoli, C. G.; Senturk, E.; Calcutt, M. W.; Rizvi, N. A., A phase I trial of SJG-136 (NSC#694501) in advanced solid tumors. *Cancer chemotherapy and pharmacology* **2010,** *65* (5), 833-8.
141. Suh, D.; Chaires, J. B., Criteria for the mode of binding of DNA binding agents. *Bioorganic & medicinal chemistry* **1995,** *3* (6), 723-8.
142. Wu, Y. S.; Koch, K. R.; Abratt, V. R.; Klump, H. H., Intercalation into the DNA double helix and in vivo biological activity of water-soluble planar [Pt(diimine)(N,N-dihydroxyethyl-N'-benzoylthioureato)](+) Cl- complexes: A study of their thermal stability, their CD spectra and their gel mobility. *Arch Biochem Biophys* **2005,** *440* (1), 28-37.
143. Martinez, R.; Chacon-Garcia, L., The search of DNA-intercalators as antitumoral drugs: what it worked and what did not work. *Current medicinal chemistry* **2005,** *12* (2), 127-51.
144. Dabrowiak, J. C., Sequence specificity of drug-DNA interactions. *Life sciences* **1983,** *32* (26), 2915-31.
145. Lane, M. J.; Dabrowiak, J. C.; Vournakis, J. N., Sequence specificity of actinomycin D and Netropsin binding to pBR322 DNA analyzed by protection from DNase I. *Proceedings of the National Academy of Sciences of the United States of America* **1983,** *80* (11), 3260-4.
146. Turner, P. R.; Ferguson, L. R.; Denny, W. A., Binding of polybenzamides to DNA: studies by DNase I and chlorambucil interference footprinting and comparison with Hoechst 33258. *Anticancer Drug Des* **1998,** *13* (8), 941-54.
147. (a) Singh, N. P.; McCoy, M. T.; Tice, R. R.; Schneider, E. L., A simple technique for quantitation of low levels of DNA damage in individual cells. *Exp Cell Res* **1988,** *175* (1), 184-91; (b) Olive, P. L.; Banath, J. P., The comet assay: a method to measure DNA damage in individual cells. *Nature protocols* **2006,** *1* (1), 23-9.
148. Hartley, J. A.; Berardini, M. D.; Souhami, R. L., An agarose gel method for the determination of DNA interstrand crosslinking applicable to the measurement of the rate of total and "second-arm" crosslink reactions. *Anal Biochem* **1991,** *193* (1), 131-4.
149. Hartley, J. A.; Hamaguchi, A.; Suggitt, M.; Gregson, S. J.; Thurston, D. E.; Howard, P. W., DNA interstrand cross-linking and in vivo antitumor activity of the extended pyrrolo[2,1-c][1,4]benzodiazepine dimer SG2057. *Invest New Drugs* **2012,** *30* (3), 950-8.
150. Martin, C.; Ellis, T.; McGurk, C. J.; Jenkins, T. C.; Hartley, J. A.; Waring, M. J.; Thurston, D. E., Sequence-selective interaction of the minor-groove interstrand cross-linking agent SJG-136 with naked and cellular DNA: footprinting and enzyme inhibition studies. *Biochemistry* **2005,** *44* (11), 4135-47.
151. Alley, M. C.; Hollingshead, M. G.; Pacula-Cox, C. M.; Waud, W. R.; Hartley, J. A.; Howard, P. W.; Gregson, S. J.; Thurston, D. E.; Sausville, E. A., SJG-136 (NSC 694501), a novel rationally designed DNA minor groove interstrand cross-linking agent with potent and broad spectrum antitumor activity: part 2: efficacy evaluations. *Cancer Res* **2004,** *64* (18), 6700-6.
152. Ledermann, J. A.; Gabra, H.; Jayson, G. C.; Spanswick, V. J.; Rustin, G. J.; Jitlal, M.; James, L. E.; Hartley, J. A., Inhibition of carboplatin-induced DNA interstrand cross-link repair by gemcitabine in patients receiving these drugs for platinum-resistant ovarian cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research* **2010,** *16* (19), 4899-905.
153. Spanswick, V. J.; Hartley, J. M.; Ward, T. H.; Hartley, J. A., Measurement of drug-induced DNA interstrand crosslinking using the single-cell gel electrophoresis (comet) assay. *Methods in molecular medicine* **1999,** *28*, 143-54.

154. Furuta, T.; Takemura, H.; Liao, Z. Y.; Aune, G. J.; Redon, C.; Sedelnikova, O. A.; Pilch, D. R.; Rogakou, E. P.; Celeste, A.; Chen, H. T.; Nussenzweig, A.; Aladjem, M. I.; Bonner, W. M.; Pommier, Y., Phosphorylation of histone H2AX and activation of Mre11, Rad50, and Nbs1 in response to replication-dependent DNA double-strand breaks induced by mammalian DNA topoisomerase I cleavage complexes. *The Journal of biological chemistry* **2003,** *278* (22), 20303-12.
155. Wu, J.; Clingen, P. H.; Spanswick, V. J.; Mellinas-Gomez, M.; Meyer, T.; Puzanov, I.; Jodrell, D.; Hochhauser, D.; Hartley, J. A., gamma-H2AX foci formation as a pharmacodynamic marker of DNA damage produced by DNA cross-linking agents: results from 2 phase I clinical trials of SJG-136 (SG2000). *Clinical cancer research : an official journal of the American Association for Cancer Research* **2013,** *19* (3), 721-30.
156. Arnould, S.; Spanswick, V. J.; Macpherson, J. S.; Hartley, J. A.; Thurston, D. E.; Jodrell, D. I.; Guichard, S. M., Time-dependent cytotoxicity induced by SJG-136 (NSC 694501): influence of the rate of interstrand cross-link formation on DNA damage signaling. *Mol Cancer Ther* **2006,** *5* (6), 1602-9.
157. Narayanaswamy, M.; Griffiths, W. J.; Howard, P. W.; Thurston, D. E., An assay combining high-performance liquid chromatography and mass spectrometry to measure DNA interstrand cross-linking efficiency in oligonucleotides of varying sequences. *Analytical biochemistry* **2008,** *374* (1), 173-81.
158. Hartley, J. A.; Spanswick, V. J.; Brooks, N.; Clingen, P. H.; McHugh, P. J.; Hochhauser, D.; Pedley, R. B.; Kelland, L. R.; Alley, M. C.; Schultz, R.; Hollingshead, M. G.; Schweikart, K. M.; Tomaszewski, J. E.; Sausville, E. A.; Gregson, S. J.; Howard, P. W.; Thurston, D. E., SJG-136 (NSC 694501), a novel rationally designed DNA minor groove interstrand cross-linking agent with potent and broad spectrum antitumor activity: part 1: cellular pharmacology, in vitro and initial in vivo antitumor activity. *Cancer research* **2004,** *64* (18), 6693-9.
159. Paull, K. D.; Shoemaker, R. H.; Hodes, L.; Monks, A.; Scudiero, D. A.; Rubinstein, L.; Plowman, J.; Boyd, M. R., Display and analysis of patterns of differential activity of drugs against human tumor cell lines: development of mean graph and COMPARE algorithm. *Journal of the National Cancer Institute* **1989,** *81* (14), 1088-92.
160. Hollingshead, M. G.; Alley, M. C.; Camalier, R. F.; Abbott, B. J.; Mayo, J. G.; Malspeis, L.; Grever, M. R., In vivo cultivation of tumor cells in hollow fibers. *Life sciences* **1995,** *57* (2), 131-41.
161. Alley, M. C.; Scudiero, D. A.; Monks, A.; Hursey, M. L.; Czerwinski, M. J.; Fine, D. L.; Abbott, B. J.; Mayo, J. G.; Shoemaker, R. H.; Boyd, M. R., Feasibility of drug screening with panels of human tumor cell lines using a microculture tetrazolium assay. *Cancer Res* **1988,** *48* (3), 589-601.
162. Monks, A.; Scudiero, D. A.; Johnson, G. S.; Paull, K. D.; Sausville, E. A., The NCI anti-cancer drug screen: a smart screen to identify effectors of novel targets. *Anticancer Drug Des* **1997,** *12* (7), 533-41.
163. Wilkinson, G. P.; Taylor, J. P.; Shnyder, S.; Cooper, P.; Howard, P. W.; Thurston, D. E.; Jenkins, T. C.; Loadman, P. M., Preliminary pharmacokinetic and bioanalytical studies of SJG-136 (NSC 694501), a sequence-selective pyrrolobenzodiazepine dimer DNA-cross-linking agent. *Invest New Drugs* **2004,** *22* (3), 231-40.
164. Reid, J. M.; Buhrow, S. A.; Kuffel, M. J.; Jia, L.; Spanswick, V. J.; Hartley, J. A.; Thurston, D. E.; Tomaszewski, J. E.; Ames, M. M., Pharmacokinetics, pharmacodynamics and metabolism of the dimeric pyrrolobenzodiazepine SJG-136 in rats. *Cancer Chemother Pharmacol* **2011,** *68* (3), 777-86.
165. Webley, S. D.; Francis, R. J.; Pedley, R. B.; Sharma, S. K.; Begent, R. H.; Hartley, J. A.; Hochhauser, D., Measurement of the critical DNA lesions produced by antibody-directed enzyme prodrug therapy (ADEPT) in vitro, in vivo and in clinical material. *British journal of cancer* **2001,** *84* (12), 1671-6.

166. Parchment, R. E.; Volpe, D. A.; LoRusso, P. M.; Erickson-Miller, C. L.; Murphy, M. J., Jr.; Grieshaber, C. K., In vivo-in vitro correlation of myelotoxicity of 9-methoxypyrazoloacridine (NSC-366140, PD115934) to myeloid and erythroid hematopoietic progenitors from human, murine, and canine marrow. *Journal of the National Cancer Institute* **1994,** *86* (4), 273-80.
167. Aguilar, M. I., HPLC of peptides and proteins: basic theory and methodology. *Methods in molecular biology* **2004,** *251*, 3-8.
168. (a) Nawrocki, J., Silica Surface Controversies, Strong Adsorption Sites, Their Blockage and Removal .2. *Chromatographia* **1991,** *31* (3-4), 193-205; (b) Nawrocki, J., Silica Surface Controversies, Strong Adsorption Sites, Their Blockage and Removal .1. *Chromatographia* **1991,** *31* (3-4), 177-192.
169. (a) Sentell, K. B.; Dorsey, J. G., Retention mechanisms in reversed-phase liquid chromatography. Stationary-phase bonding density and partitioning. *Analytical chemistry* **1989,** *61* (9), 930-4; (b) Sentell, K. B.; Dorsey, J. G., Retention mechanisms in reversed-phase chromatography. Stationary phase bonding density and solute selectivity. *Journal of chromatography* **1989,** *461*, 193-207.
170. Dorsey, J. G.; Cooper, W. T., Retention Mechanisms of Bonded-Phase Liquid-Chromatography. *Analytical chemistry* **1994,** *66* (17), A857-A867.
171. Dill, K. A.; Naghizadeh, J.; Marqusee, J. A., Chain molecules at high densities at interfaces. *Annual review of physical chemistry* **1988,** *39*, 425-61.
172. Oefner, P. J.; Huber, C. G., A decade of high-resolution liquid chromatography of nucleic acids on styrene-divinylbenzene copolymers. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences* **2002,** *782* (1-2), 27-55.
173. Castleberry, C. M.; Chou, C. W.; Limbach, P. A., Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry of oligonucleotides. *Current protocols in nucleic acid chemistry / edited by Serge L. Beaucage ... [et al.]* **2008,** *Chapter 10*, Unit 10 1.
174. Tang, K.; Allman, S. L.; Jones, R. B.; Chen, C. H.; Araghi, S., Laser mass spectrometry of oligonucleotides with isomer matrices. *Rapid communications in mass spectrometry : RCM* **1993,** *7* (6), 435-9.
175. Beavis, R. C.; Chait, B. T., Factors affecting the ultraviolet laser desorption of proteins. *Rapid communications in mass spectrometry : RCM* **1989,** *3* (7), 233-7.
176. (a) Wu, K. J.; Steding, A.; Becker, C. H., Matrix-assisted laser desorption time-of-flight mass spectrometry of oligonucleotides using 3-hydroxypicolinic acid as an ultraviolet-sensitive matrix. *Rapid communications in mass spectrometry : RCM* **1993,** *7* (2), 142-6; (b) Wu, K. J.; Shaler, T. A.; Becker, C. H., Time-of-flight mass spectrometry of underivatized single-stranded DNA oligomers by matrix-assisted laser desorption. *Analytical chemistry* **1994,** *66* (10), 1637-45.
177. Nordhoff, E.; Cramer, R.; Karas, M.; Hillenkamp, F.; Kirpekar, F.; Kristiansen, K.; Roepstorff, P., Ion stability of nucleic acids in infrared matrix-assisted laser desorption/ionization mass spectrometry. *Nucleic acids research* **1993,** *21* (15), 3347-57.
178. (a) Lewis, D. G.; Johnson, W. C., Jr., Circular dichroism of DNA in the vacuum ultraviolet. *Journal of molecular biology* **1974,** *86* (1), 91-6; (b) Holm, A. I.; Nielsen, L. M.; Hoffmann, S. V.; Nielsen, S. B., Vacuum-ultraviolet circular dichroism spectroscopy of DNA: a valuable tool to elucidate topology and electronic coupling in DNA. *Physical chemistry chemical physics : PCCP* **2010,** *12* (33), 9581-96; (c) Kypr, J.; Kejnovska, I.; Renciuk, D.; Vorlickova, M., Circular dichroism and conformational polymorphism of DNA. *Nucleic acids research* **2009,** *37* (6), 1713-25.
179. Woody, R. W., Circular-Dichroism. *Method Enzymol* **1995,** *246*, 34-71.
180. Ho, P. S.; Zhou, G. W.; Clark, L. B., Polarized electronic spectra of Z-DNA single crystals. *Biopolymers* **1990,** *30* (1-2), 151-63.
181. (a) Gray, D. M.; Liu, J. J.; Ratliff, R. L.; Allen, F. S., Sequence Dependence of the Circular-Dichroism of Synthetic Double-Stranded Rnas. *Biopolymers* **1981,** *20* (7), 1337-1382; (b) Cox, R.

A.; Hirst, W.; Godwin, E.; Kaiser, I., Circular-Dichroism of Ribosomal Ribonucleic-Acids. *Biochemical Journal* **1976,** *155* (2), 279-291.
182. Kankia, B. I.; Buckin, V.; Bloomfield, V. A., Hexamminecobalt(III)-induced condensation of calf thymus DNA: circular dichroism and hydration measurements. *Nucleic acids research* **2001,** *29* (13), 2795-801.
183. Maestre, M. F., Circular dichroism of DNA films: reversibility studies. *Journal of molecular biology* **1970,** *52* (3), 543-56.
184. (a) Forster, T., Energiewanderung Und Fluoreszenz. *Naturwissenschaften* **1946,** *33* (6), 166-175; (b) Forster, T., *Zwischenmolekulare Energiewanderung Und Fluoreszenz. *Ann Phys-Berlin* **1948,** *2* (1-2), 55-75.
185. Jonsson, T.; Waldburger, C. D.; Sauer, R. T., Nonlinear free energy relationships in Arc repressor unfolding imply the existence of unstable, native-like folding intermediates. *Biochemistry* **1996,** *35* (15), 4795-802.
186. Clegg, R. M.; Murchie, A. I.; Lilley, D. M., The solution structure of the four-way DNA junction at low-salt conditions: a fluorescence resonance energy transfer analysis. *Biophysical journal* **1994,** *66* (1), 99-109.
187. Parkhurst, K. M.; Parkhurst, L. J., Kinetic studies by fluorescence resonance energy transfer employing a double-labeled oligonucleotide: hybridization to the oligonucleotide complement and to single-stranded DNA. *Biochemistry* **1995,** *34* (1), 285-92.
188. Moore, G. E., Cramming More Components Onto Integrated Circuits. *Proceedings of the IEEE* **1998,** *86* (1), 82-85.
189. Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y.; Wriggers, W., Atomic-level characterization of the structural dynamics of proteins. *Science* **2010,** *330* (6002), 341-6.
190. Shaw, D. E.; Deneroff, M. M.; Dror, R. O.; Kuskin, J. S.; Larson, R. H.; Salmon, J. K.; Young, C.; Batson, B.; Bowers, K. J.; Chao, J. C.; Eastwood, M. P.; Gagliardo, J.; Grossman, J. P.; Ho, C. R.; Ierardi, D. J.; Istv; #225; Kolossv, n.; ry; Klepeis, J. L.; Layman, T.; McLeavey, C.; Moraes, M. A.; Mueller, R.; Priest, E. C.; Shan, Y.; Spengler, J.; Theobald, M.; Towles, B.; Wang, S. C., Anton, a special-purpose machine for molecular dynamics simulation. In *Proceedings of the 34th annual international symposium on Computer architecture*, ACM: San Diego, California, USA, 2007; pp 1-12.
191. (a) Brucoli, F.; Hawkins, R. M.; James, C. H.; Jackson, P. J.; Wells, G.; Jenkins, T. C.; Ellis, T.; Kotecha, M.; Hochhauser, D.; Hartley, J. A.; Howard, P. W.; Thurston, D. E., An Extended Pyrrolobenzodiazepine-Polyamide Conjugate with Selectivity for a DNA Sequence Containing the ICB2 Transcription Factor Binding Site. *J Med Chem* **2013,** *56* (16), 6339-51; (b) Mantaj, J., Jackson, P. J. M., Rahman, K. M., Thurston, D. E. In *Interaction of SJG-136 with cognate sequences of oncogenic transcription factors*, American Association of Cancer Research Annual Meeting, Washington, Washington, 2013; (c) Jackson, P. J. M.; James, C. H.; Jenkins, T. C.; Rahman, K. M.; Thurston, D. E., Computational Studies Support the Role of the C7-Sibirosamine Sugar of the Pyrrolobenzodiazepine (PBD) Sibiromycin in Transcription Factor Inhibition. *ACS Chemical Biology* **2014**; (d) Jackson, P. J. M.; Rahman, K. M.; Thurston, D. E., Abstract 3696: Molecular dynamics simulations of C7/C7'-linked pyrrolobenzodiazepine (PBD) dimers. *Cancer research* **2015,** *75* (15 Supplement), 3696.
192. Saiki, R. K.; Scharf, S.; Faloona, F.; Mullis, K. B.; Horn, G. T.; Erlich, H. A.; Arnheim, N., Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **1985,** *230* (4732), 1350-4.
193. Mullis, K. B., The unusual origin of the polymerase chain reaction. *Scientific American* **1990,** *262* (4), 56-61, 64-5.
194. Wong, M. L.; Medrano, J. F., Real-time PCR for mRNA quantitation. *BioTechniques* **2005,** *39* (1), 75-85.

195. Valasek, M. A.; Repa, J. J., The power of real-time PCR. *Advances in physiology education* **2005,** *29* (3), 151-9.
196. Kubista, M.; Andrade, J. M.; Bengtsson, M.; Forootan, A.; Jonak, J.; Lind, K.; Sindelka, R.; Sjoback, R.; Sjogreen, B.; Strombom, L.; Stahlberg, A.; Zoric, N., The real-time polymerase chain reaction. *Molecular aspects of medicine* **2006,** *27* (2-3), 95-125.
197. Wittwer, C. T.; Herrmann, M. G.; Moss, A. A.; Rasmussen, R. P., Continuous fluorescence monitoring of rapid cycle DNA amplification. *BioTechniques* **1997,** *22* (1), 130-1, 134-8.
198. Mackay, I. M., Real-time PCR in the microbiology laboratory. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* **2004,** *10* (3), 190-212.
199. Gibson, U. E.; Heid, C. A.; Williams, P. M., A novel method for real time quantitative RT-PCR. *Genome research* **1996,** *6* (10), 995-1001.
200. Rahman, K. M.; Jackson, P. J. M.; James, C. H.; Basu, B. P.; Hartley, J. A.; de la Fuente, M.; Schatzlein, A.; Robson, M.; Pedley, R. B.; Pepper, C.; Fox, K. R.; Howard, P. W.; Thurston, D. E., GC-Targeted C8-Linked Pyrrolobenzodiazepine–Biaryl Conjugates with Femtomolar in Vitro Cytotoxicity and in Vivo Antitumor Activity in Mouse Models. *J. Med. Chem.* **2013,** *56* (7), 2911-2935.
201. Mantaj, J.; Jackson, P. J.; Karu, K.; Rahman, K. M.; Thurston, D. E., Covalent Bonding of Pyrrolobenzodiazepines (PBDs) to Terminal Guanine Residues within Duplex and Hairpin DNA Fragments. *PloS one* **2016,** *11* (4), e0152303.
202. Rahman, K. M.; Mussa, V.; Narayanaswamy, M.; James, C. H.; Howard, P. W.; Thurston, D. E., Observation of a dynamic equilibrium between DNA hairpin and duplex forms of covalent adducts of a minor groove binding agent. *Chem Commun (Camb)* **2009,** (2), 227-9.
203. Puvvada, M. S.; Hartley, J. A.; Jenkins, T. C.; Thurston, D. E., A quantitative assay to measure the relative DNA-binding affinity of pyrrolo[2,1-c] [1,4]benzodiazepine (PBD) antitumour antibiotics based on the inhibition of restriction endonuclease BamHI. *Nucleic acids research* **1993,** *21* (16), 3671-5.
204. (a) Chang, Y. M.; Chen, C. K.; Hou, M. H., Conformational Changes in DNA upon Ligand Binding Monitored by Circular Dichroism. *International journal of molecular sciences* **2012,** *13* (3), 3394-413; (b) Kaushik, M.; Kukreti, R.; Grover, D.; Brahmachari, S. K.; Kukreti, S., Hairpin-duplex equilibrium reflected in the A-->B transition in an undecamer quasi-palindrome present in the locus control region of the human beta-globin gene cluster. *Nucleic acids research* **2003,** *31* (23), 6904-15.
205. Rahman, K. M.; Thompson, A. S.; James, C. H.; Narayanaswamy, M.; Thurston, D. E., The Pyrrolobenzodiazepine Dimer SJG-136 Forms Sequence-Dependent Intrastrand DNA Cross-Links and Monoalkylated Adducts in Addition to Interstrand Cross-Links. *Journal of the American Chemical Society* **2009,** *131* (38), 13756-13766.
206. Thurston, D. E.; Vassoler, H.; Jackson, P. J. M.; James, C. H.; Rahman, K. M., Effect of hairpin loop structure on reactivity, sequence preference and adduct orientation of a DNA-interactive pyrrolo[2,1-c][1,4]benzodiazepine (PBD) antitumour agent. *Organic & biomolecular chemistry* **2015,** *13* (13), 4031-4040.
207. Martin, C.; Ellis, T.; McGurk, C. J.; Jenkins, T. C.; Hartley, J. A.; Waring, M. J.; Thurston, D. E., Sequence-Selective Interaction of the Minor-Groove Interstrand Cross-Linking Agent SJG-136 with Naked and Cellular DNA: Footprinting and Enzyme Inhibition Studies†. *Biochemistry* **2005,** *44* (11), 4135-4147.
208. Thurston, D. E.; Vassoler, H.; Jackson, P. J.; James, C. H.; Rahman, K. M., Effect of Hairpin Loop Structure on Reactivity, Sequence Preference and Adduct Orientation of a DNA-Interactive Pyrrolo[2,1-*c*][1,4]benzodiazepine (PBD) Antitumour Agent. *Organic & biomolecular chemistry* **2015,** *13* (13), 4031-40.

209. Hopton, S. R.; Thompson, A. S., Nuclear magnetic resonance solution structures of inter- and intrastrand adducts of DNA cross-linker SJG-136. *Biochemistry* **2011,** *50* (21), 4720-32.
210. (a) Rahman, K. M.; James, C. H.; Bui, T. T.; Drake, A. F.; Thurston, D. E., Observation of a single-stranded DNA/pyrrolobenzodiazepine adduct. *J Am Chem Soc* **2011,** *133* (48), 19376-85; (b) Rahman, K. M.; James, C. H.; Thurston, D. E., Effect of Base Sequence on the DNA Cross-Linking Properties of Pyrrolobenzodiazepine (PBD) Dimers. *Nucleic acids research* **2011,** *39* (13), 5800-5812; (c) Rahman, K. M.; Jackson, P. J.; James, C. H.; Basu, B. P.; Hartley, J. A.; de la Fuente, M.; Schatzlein, A.; Robson, M.; Pedley, R. B.; Pepper, C., GC-Targeted C8-Linked Pyrrolobenzodiazepine–Biaryl Conjugates with Femtomolar in Vitro Cytotoxicity and in Vivo Antitumor Activity in Mouse Models. *J. Med. Chem.* **2013,** *56* (7), 2911-2935.
211. (a) Glinghammar, B.; Holmberg, K.; Rafter, J., Effects of colonic lumenal components on AP-1-dependent gene transcription in cultured human colon carcinoma cells. *Carcinogenesis* **1999,** *20* (6), 969-76; (b) Matheson, H.; Branting, C.; Rafter, I.; Okret, S.; Rafter, J., Increased c-fos mRNA and binding to the AP-1 recognition sequence accompanies the proliferative response to deoxycholate of HT29 cells. *Carcinogenesis* **1996,** *17* (3), 421-6.
212. (a) Snyder, M.; Huang, X. Y.; Zhang, J. J., Signal transducers and activators of transcription 3 (STAT3) directly regulates cytokine-induced fascin expression and is required for breast cancer cell migration. *The Journal of biological chemistry* **2011,** *286* (45), 38886-93; (b) Ling, X.; Arlinghaus, R. B., Knockdown of STAT3 expression by RNA interference inhibits the induction of breast tumors in immunocompetent mice. *Cancer research* **2005,** *65* (7), 2532-6.
213. Berg, J. M.; Tymoczko, J. L.; Stryer, L.; Stryer, L., *Biochemistry*. 5th ed.; W.H. Freeman: New York, 2002.
214. (a) Alberts, B., *Molecular biology of the cell*. 4th ed.; Garland Science: New York, 2002; p xxxiv, 1548 p; (b) Lodish, H. F., *Molecular cell biology*. 5th ed.; W.H. Freeman and Company: New York, 2003; p xxxiii, 973 , 79 p.
215. Chen, M.; Manley, J. L., Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nature reviews. Molecular cell biology* **2009,** *10* (11), 741-54.
216. (a) Merino, E.; Yanofsky, C., Transcription attenuation: a highly conserved regulatory strategy used by bacteria. *Trends in genetics : TIG* **2005,** *21* (5), 260-4; (b) Yanofsky, C., Transcription attenuation. *The Journal of biological chemistry* **1988,** *263* (2), 609-12; (c) Yanofsky, C., Transcription attenuation: once viewed as a novel regulatory strategy. *Journal of bacteriology* **2000,** *182* (1), 1-8.
217. (a) Gott, J. M.; Emeson, R. B., Functions and mechanisms of RNA editing. *Annual review of genetics* **2000,** *34*, 499-531; (b) Brennicke, A.; Marchfelder, A.; Binder, S., RNA editing. *FEMS microbiology reviews* **1999,** *23* (3), 297-316.
218. Wek, R. C.; Jiang, H. Y.; Anthony, T. G., Coping with stress: eIF2 kinases and translational control. *Biochemical Society transactions* **2006,** *34* (Pt 1), 7-11.
219. (a) Fernandez, J.; Yaman, I.; Mishra, R.; Merrick, W. C.; Snider, M. D.; Lamers, W. H.; Hatzoglou, M., Internal ribosome entry site-mediated translation of a mammalian mRNA is regulated by amino acid availability. *The Journal of biological chemistry* **2001,** *276* (15), 12285-91; (b) Martinez-Salas, E.; Pineiro, D.; Fernandez, N., Alternative Mechanisms to Initiate Translation in Eukaryotic mRNAs. *Comparative and functional genomics* **2012,** *2012*, 391546.
220. (a) Sheets, M. D.; Ogg, S. C.; Wickens, M. P., Point mutations in AAUAAA and the poly (A) addition site: effects on the accuracy and efficiency of cleavage and polyadenylation in vitro. *Nucleic acids research* **1990,** *18* (19), 5799-805; (b) Korner, C. G.; Wahle, E., Poly(A) tail shortening by a mammalian poly(A)-specific 3'-exoribonuclease. *The Journal of biological chemistry* **1997,** *272* (16), 10448-56.
221. (a) Brogna, S.; Wen, J., Nonsense-mediated mRNA decay (NMD) mechanisms. *Nature structural & molecular biology* **2009,** *16* (2), 107-13; (b) Schweingruber, C.; Rufener, S. C.;

Zund, D.; Yamashita, A.; Muhlemann, O., Nonsense-mediated mRNA decay - mechanisms of substrate mRNA recognition and degradation in mammalian cells. *Biochimica et biophysica acta* **2013,** *1829* (6-7), 612-23.
222. Brinkley, B. R.; Beall, P. T.; Wible, L. J.; Mace, M. L.; Turner, D. S.; Cailleau, R. M., Variations in cell form and cytoskeleton in human breast carcinoma cells in vitro. *Cancer research* **1980,** *40* (9), 3118-29.
223. Cailleau, R.; Olive, M.; Cruciger, Q. V., Long-term human breast carcinoma cell lines of metastatic origin: preliminary characterization. *In vitro* **1978,** *14* (11), 911-5.
224. Siciliano, M. J.; Barker, P. E.; Cailleau, R., Mutually exclusive genetic signatures of human breast tumor cell lines with a common chromosomal marker. *Cancer research* **1979,** *39* (3), 919-22.
225. Barbieri, I.; Pensa, S.; Pannellini, T.; Quaglino, E.; Maritano, D.; Demaria, M.; Voster, A.; Turkson, J.; Cavallo, F.; Watson, C. J.; Provero, P.; Musiani, P.; Poli, V., Constitutively active Stat3 enhances neu-mediated migration and metastasis in mammary tumors via upregulation of Cten. *Cancer research* **2010,** *70* (6), 2558-67.
226. (a) Cohen, E.; Ophir, I.; Shaul, Y. B., Induced differentiation in HT29, a human colon adenocarcinoma cell line. *Journal of cell science* **1999,** *112 (Pt 16)*, 2657-66; (b) Hekmati, M.; Polak-Charcon, S.; Ben-Shaul, Y., A morphological study of a human adenocarcinoma cell line (HT29) differentiating in culture. Similarities to intestinal embryonic development. *Cell differentiation and development : the official journal of the International Society of Developmental Biologists* **1990,** *31* (3), 207-18.
227. Bishnupuri, K. S.; Luo, Q.; Murmu, N.; Houchen, C. W.; Anant, S.; Dieckgraefe, B. K., Reg IV activates the epidermal growth factor receptor/Akt/AP-1 signaling pathway in colon adenocarcinomas. *Gastroenterology* **2006,** *130* (1), 137-49.
228. Yip, K. W.; Reed, J. C., Bcl-2 family proteins and cancer. *Oncogene* **2008,** *27* (50), 6398-406.
229. (a) Cory, S.; Huang, D. C.; Adams, J. M., The Bcl-2 family: roles in cell survival and oncogenesis. *Oncogene* **2003,** *22* (53), 8590-607; (b) Levine, B.; Kroemer, G., Autophagy in the pathogenesis of disease. *Cell* **2008,** *132* (1), 27-42; (c) Reed, J. C., Bcl-2-family proteins and hematologic malignancies: history and future prospects. *Blood* **2008,** *111* (7), 3322-30.
230. Tsujimoto, Y.; Cossman, J.; Jaffe, E.; Croce, C. M., Involvement of the bcl-2 gene in human follicular lymphoma. *Science* **1985,** *228* (4706), 1440-3.
231. (a) Cimmino, A.; Calin, G. A.; Fabbri, M.; Iorio, M. V.; Ferracin, M.; Shimizu, M.; Wojcik, S. E.; Aqeilan, R. I.; Zupo, S.; Dono, M.; Rassenti, L.; Alder, H.; Volinia, S.; Liu, C. G.; Kipps, T. J.; Negrini, M.; Croce, C. M., miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proceedings of the National Academy of Sciences of the United States of America* **2005,** *102* (39), 13944-9; (b) Hanada, M.; Delia, D.; Aiello, A.; Stadtmauer, E.; Reed, J. C., bcl-2 gene hypomethylation and high-level expression in B-cell chronic lymphocytic leukemia. *Blood* **1993,** *82* (6), 1820-8.
232. Vermeulen, K.; Van Bockstaele, D. R.; Berneman, Z. N., The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer. *Cell proliferation* **2003,** *36* (3), 131-49.
233. (a) Hall, M.; Peters, G., Genetic alterations of cyclins, cyclin-dependent kinases, and Cdk inhibitors in human cancer. *Advances in cancer research* **1996,** *68*, 67-108; (b) Gillett, C.; Smith, P.; Gregory, W.; Richards, M.; Millis, R.; Peters, G.; Barnes, D., Cyclin D1 and prognosis in human breast cancer. *International journal of cancer. Journal international du cancer* **1996,** *69* (2), 92-9; (c) Motokura, T.; Arnold, A., Cyclins and oncogenesis. *Biochimica et biophysica acta* **1993,** *1155* (1), 63-78; (d) Musgrove, E. A., Cyclins: roles in mitogenic signaling and oncogenic transformation. *Growth factors* **2006,** *24* (1), 13-9; (e) Sutherland, R. L.; Musgrove, E. A., Cyclin D1 and mammary carcinoma: new insights from transgenic mouse models. *Breast cancer research : BCR* **2002,** *4* (1), 14-7; (f) Weinstat-Saslow, D.; Merino, M. J.; Manrow, R. E.; Lawrence, J. A.; Bluth, R. F.; Wittenbel, K. D.; Simpson, J. F.; Page, D. L.; Steeg, P. S.,

Overexpression of cyclin D mRNA distinguishes invasive and in situ breast carcinomas from non-malignant lesions. *Nature medicine* **1995,** *1* (12), 1257-60; (g) Ewen, M. E.; Lamb, J., The activities of cyclin D1 that drive tumorigenesis. *Trends in molecular medicine* **2004,** *10* (4), 158-62.
234. Seethala, R. R.; Gooding, W. E.; Handler, P. N.; Collins, B.; Zhang, Q.; Siegfried, J. M.; Grandis, J. R., Immunohistochemical analysis of phosphotyrosine signal transducer and activator of transcription 3 and epidermal growth factor receptor autocrine signaling pathways in head and neck cancers and metastatic lymph nodes. *Clinical cancer research : an official journal of the American Association for Cancer Research* **2008,** *14* (5), 1303-9.
235. Stephanou, A.; Latchman, D. S., STAT-1: a novel regulator of apoptosis. *International journal of experimental pathology* **2003,** *84* (6), 239-44.
236. (a) Ramana, C. V.; Chatterjee-Kishore, M.; Nguyen, H.; Stark, G. R., Complex roles of Stat1 in regulating gene expression. *Oncogene* **2000,** *19* (21), 2619-27; (b) Avalle, L.; Pensa, S.; Regis, G.; Novelli, F.; Poli, V., STAT1 and STAT3 in tumorigenesis: A matter of balance. *Jak-Stat* **2012,** *1* (2), 65-72; (c) Zimmerer, J. M.; Lesinski, G. B.; Radmacher, M. D.; Ruppert, A.; Carson, W. E., 3rd, STAT1-dependent and STAT1-independent gene expression in murine immune cells following stimulation with interferon-alpha. *Cancer immunology, immunotherapy : CII* **2007,** *56* (11), 1845-52.
237. (a) Barre, B.; Avril, S.; Coqueret, O., Opposite regulation of myc and p21waf1 transcription by STAT3 proteins. *The Journal of biological chemistry* **2003,** *278* (5), 2990-6; (b) Kominsky, S. L.; Hobeika, A. C.; Lake, F. A.; Torres, B. A.; Johnson, H. M., Down-regulation of neu/HER-2 by interferon-gamma in prostate cancer cells. *Cancer research* **2000,** *60* (14), 3904-8.
238. Machesky, L. M.; Li, A., Fascin: Invasive filopodia promoting metastasis. *Communicative & integrative biology* **2010,** *3* (3), 263-70.
239. Khwaja, A.; Rodriguez-Viciana, P.; Wennstrom, S.; Warne, P. H.; Downward, J., Matrix adhesion and Ras transformation both activate a phosphoinositide 3-OH kinase and protein kinase B/Akt cellular survival pathway. *The EMBO journal* **1997,** *16* (10), 2783-93.
240. Jiang, P.; Enomoto, A.; Takahashi, M., Cell biology of the movement of breast cancer cells: intracellular signalling and the actin cytoskeleton. *Cancer letters* **2009,** *284* (2), 122-30.
241. Aksoy, S.; Szumlanski, C. L.; Weinshilboum, R. M., Human liver nicotinamide N-methyltransferase. cDNA cloning, expression, and biochemical characterization. *The Journal of biological chemistry* **1994,** *269* (20), 14835-40.
242. Xu, J.; Moatamed, F.; Caldwell, J. S.; Walker, J. R.; Kraiem, Z.; Taki, K.; Brent, G. A.; Hershman, J. M., Enhanced expression of nicotinamide N-methyltransferase in human papillary thyroid carcinoma cells. *The Journal of clinical endocrinology and metabolism* **2003,** *88* (10), 4990-6.
243. Yamada, K.; Miyazaki, T.; Hara, N.; Tsuchiya, M., Interferon-gamma elevates nicotinamide N-methyltransferase activity and nicotinamide level in human glioma cells. *Journal of nutritional science and vitaminology* **2010,** *56* (2), 83-6.
244. Lim, B. H.; Cho, B. I.; Kim, Y. N.; Kim, J. W.; Park, S. T.; Lee, C. W., Overexpression of nicotinamide N-methyltransferase in gastric cancer tissues and its potential post-translational modification. *Experimental & molecular medicine* **2006,** *38* (5), 455-65.
245. Yao, M.; Tabuchi, H.; Nagashima, Y.; Baba, M.; Nakaigawa, N.; Ishiguro, H.; Hamada, K.; Inayama, Y.; Kishida, T.; Hattori, K.; Yamada-Okabe, H.; Kubota, Y., Gene expression analysis of renal carcinoma: adipose differentiation-related protein as a potential diagnostic and prognostic biomarker for clear-cell renal carcinoma. *The Journal of pathology* **2005,** *205* (3), 377-87.
246. Tomida, M.; Mikami, I.; Takeuchi, S.; Nishimura, H.; Akiyama, H., Serum levels of nicotinamide N-methyltransferase in patients with lung cancer. *Journal of cancer research and clinical oncology* **2009,** *135* (9), 1223-9.

247. Rogers, C. D.; Fukushima, N.; Sato, N.; Shi, C.; Prasad, N.; Hustinx, S. R.; Matsubayashi, H.; Canto, M.; Eshleman, J. R.; Hruban, R. H.; Goggins, M., Differentiating pancreatic lesions by microarray and QPCR analysis of pancreatic juice RNAs. *Cancer biology & therapy* **2006,** *5* (10), 1383-9.
248. Roessler, M.; Rollinger, W.; Palme, S.; Hagmann, M. L.; Berndt, P.; Engel, A. M.; Schneidinger, B.; Pfeffer, M.; Andres, H.; Karl, J.; Bodenmuller, H.; Ruschoff, J.; Henkel, T.; Rohr, G.; Rossol, S.; Rosch, W.; Langen, H.; Zolg, W.; Tacke, M., Identification of nicotinamide N-methyltransferase as a novel serum tumor marker for colorectal cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research* **2005,** *11* (18), 6550-7.
249. Tsuchiya, A.; Sakamoto, M.; Yasuda, J.; Chuma, M.; Ohta, T.; Ohki, M.; Yasugi, T.; Taketani, Y.; Hirohashi, S., Expression profiling in ovarian clear cell carcinoma: identification of hepatocyte nuclear factor-1 beta as a molecular marker and a possible molecular target for therapy of ovarian clear cell carcinoma. *The American journal of pathology* **2003,** *163* (6), 2503-12.
250. Wu, Y.; Siadaty, M. S.; Berens, M. E.; Hampton, G. M.; Theodorescu, D., Overlapping gene expression profiles of cell migration and tumor invasion in human bladder cancer identify metallothionein 1E and nicotinamide N-methyltransferase as novel regulators of cell migration. *Oncogene* **2008,** *27* (52), 6679-89.
251. Emanuelli, M.; Santarelli, A.; Sartini, D.; Ciavarella, D.; Rossi, V.; Pozzi, V.; Rubini, C.; Lo Muzio, L., Nicotinamide N-Methyltransferase upregulation correlates with tumour differentiation in oral squamous cell carcinoma. *Histology and histopathology* **2010,** *25* (1), 15-20.
252. Thomas, M. G.; Saldanha, M.; Mistry, R. J.; Dexter, D. T.; Ramsden, D. B.; Parsons, R. B., Nicotinamide N-methyltransferase expression in SH-SY5Y neuroblastoma and N27 mesencephalic neurones induces changes in cell morphology via ephrin-B2 and Akt signalling. *Cell death & disease* **2013,** *4*, e669.
253. Tang, S. W.; Yang, T. C.; Lin, W. C.; Chang, W. H.; Wang, C. C.; Lai, M. K.; Lin, J. Y., Nicotinamide N-methyltransferase induces cellular invasion through activating matrix metalloproteinase-2 expression in clear cell renal cell carcinoma cells. *Carcinogenesis* **2011,** *32* (2), 138-45.
254. Ferrara, N.; Houck, K.; Jakeman, L.; Leung, D. W., Molecular and biological properties of the vascular endothelial growth factor family of proteins. *Endocrine reviews* **1992,** *13* (1), 18-32.
255. Gerber, H. P.; Vu, T. H.; Ryan, A. M.; Kowalski, J.; Werb, Z.; Ferrara, N., VEGF couples hypertrophic cartilage remodeling, ossification and angiogenesis during endochondral bone formation. *Nature medicine* **1999,** *5* (6), 623-8.
256. Ferrara, N.; Carver-Moore, K.; Chen, H.; Dowd, M.; Lu, L.; O'Shea, K. S.; Powell-Braxton, L.; Hillan, K. J.; Moore, M. W., Heterozygous embryonic lethality induced by targeted inactivation of the VEGF gene. *Nature* **1996,** *380* (6573), 439-42.
257. Chintalgattu, V.; Nair, D. M.; Katwa, L. C., Cardiac myofibroblasts: a novel source of vascular endothelial growth factor (VEGF) and its receptors Flt-1 and KDR. *Journal of molecular and cellular cardiology* **2003,** *35* (3), 277-86.
258. Carmeliet, P., VEGF as a key mediator of angiogenesis in cancer. *Oncology* **2005,** *69 Suppl 3*, 4-10.
259. Poon, R. T.; Fan, S. T.; Wong, J., Clinical implications of circulating angiogenic factors in cancer patients. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **2001,** *19* (4), 1207-25.
260. (a) Wang, X.; Ren, Y.; Zhuang, H.; Meng, X.; Huang, S.; Li, Y.; Hehir, M.; Wang, P., Decrease of phosphorylated proto-oncogene CREB at Ser 133 site inhibits growth and metastatic activity of renal cell cancer. *Expert opinion on therapeutic targets* **2015,** *19* (7), 985-

95; (b) Gonzalez, G. A.; Montminy, M. R., Cyclic AMP stimulates somatostatin gene transcription by phosphorylation of CREB at serine 133. *Cell* **1989,** *59* (4), 675-80.
261. Abramovitch, R.; Tavor, E.; Jacob-Hirsch, J.; Zeira, E.; Amariglio, N.; Pappo, O.; Rechavi, G.; Galun, E.; Honigman, A., A pivotal role of cyclic AMP-responsive element binding protein in tumor progression. *Cancer research* **2004,** *64* (4), 1338-46.
262. (a) Shaywitz, A. J.; Greenberg, M. E., CREB: a stimulus-induced transcription factor activated by a diverse array of extracellular signals. *Annual review of biochemistry* **1999,** *68*, 821-61; (b) Mayr, B.; Montminy, M., Transcriptional regulation by the phosphorylation-dependent factor CREB. *Nature reviews. Molecular cell biology* **2001,** *2* (8), 599-609.
263. Xiao, X.; Li, B. X.; Mitton, B.; Ikeda, A.; Sakamoto, K. M., Targeting CREB for cancer therapy: friend or foe. *Current cancer drug targets* **2010,** *10* (4), 384-91.
264. (a) Tan, X.; Wang, S.; Zhu, L.; Wu, C.; Yin, B.; Zhao, J.; Yuan, J.; Qiang, B.; Peng, X., cAMP response element-binding protein promotes gliomagenesis by modulating the expression of oncogenic microRNA-23a. *Proceedings of the National Academy of Sciences of the United States of America* **2012,** *109* (39), 15805-10; (b) Rodon, L.; Gonzalez-Junca, A.; Inda Mdel, M.; Sala-Hojman, A.; Martinez-Saez, E.; Seoane, J., Active CREB1 promotes a malignant TGFbeta2 autocrine loop in glioblastoma. *Cancer discovery* **2014,** *4* (10), 1230-41.
265. (a) Chhabra, A.; Fernando, H.; Watkins, G.; Mansel, R. E.; Jiang, W. G., Expression of transcription factor CREB1 in human breast cancer and its correlation with prognosis. *Oncology reports* **2007,** *18* (4), 953-8; (b) Zhang, S.; Chen, L.; Cui, B.; Chuang, H. Y.; Yu, J.; Wang-Rodriguez, J.; Tang, L.; Chen, G.; Basak, G. W.; Kipps, T. J., ROR1 is expressed in human breast cancer and associated with enhanced tumor-cell growth. *PloS one* **2012,** *7* (3), e31127.
266. Seo, H. S.; Liu, D. D.; Bekele, B. N.; Kim, M. K.; Pisters, K.; Lippman, S. M.; Wistuba, II; Koo, J. S., Cyclic AMP response element-binding protein overexpression: a feature associated with negative prognosis in never smokers with non-small cell lung cancer. *Cancer research* **2008,** *68* (15), 6065-73.
267. Wu, D.; Zhau, H. E.; Huang, W. C.; Iqbal, S.; Habib, F. K.; Sartor, O.; Cvitanovic, L.; Marshall, F. F.; Xu, Z.; Chung, L. W., cAMP-responsive element-binding protein regulates vascular endothelial growth factor expression: implication in human prostate cancer bone metastasis. *Oncogene* **2007,** *26* (35), 5070-7.
268. Crans-Vargas, H. N.; Landaw, E. M.; Bhatia, S.; Sandusky, G.; Moore, T. B.; Sakamoto, K. M., Expression of cyclic adenosine monophosphate response-element binding protein in acute leukemia. *Blood* **2002,** *99* (7), 2617-9.
269. Bordonaro, M.; Lazarova, D. L., CREB-binding protein, p300, butyrate, and Wnt signaling in colorectal cancer. *World journal of gastroenterology : WJG* **2015,** *21* (27), 8238-48.
270. Sharrocks, A. D., The ETS-domain transcription factor family. *Nature reviews. Molecular cell biology* **2001,** *2* (11), 827-37.
271. Sharrocks, A. D.; Yang, S. H.; Galanis, A., Docking domains and substrate-specificity determination for MAP kinases. *Trends in biochemical sciences* **2000,** *25* (9), 448-53.
272. (a) Kasza, A., Signal-dependent Elk-1 target genes involved in transcript processing and cell migration. *Biochimica et biophysica acta* **2013,** *1829* (10), 1026-33; (b) Hipskind, R. A.; Rao, V. N.; Mueller, C. G.; Reddy, E. S.; Nordheim, A., Ets-related protein Elk-1 is homologous to the c-fos regulatory factor p62TCF. *Nature* **1991,** *354* (6354), 531-4.
273. Shaulian, E., AP-1 — The Jun proteins: Oncogenes or tumor suppressors in disguise? *Cellular Signalling* **2010,** *22* (6), 894-899.
274. Kappelmann, M.; Bosserhoff, A.; Kuphal, S., AP-1/c-Jun transcription factors: Regulation and function in malignant melanoma. *European Journal of Cell Biology* (0).
275. Jin, X.; Song, X.; Li, L.; Wang, Z.; Tao, Y.; Deng, L.; Tang, M.; Yi, W.; Cao, Y., Blockade of AP-1 activity by dominant-negative TAM67 can abrogate the oncogenic phenotype in latent membrane protein 1-positive human nasopharyngeal carcinoma. *Molecular carcinogenesis* **2007,** *46* (11), 901-11.

276. Blau, L.; Knirsh, R.; Ben-Dror, I.; Oren, S.; Kuphal, S.; Hau, P.; Proescholdt, M.; Bosserhoff, A. K.; Vardimon, L., Aberrant expression of c-Jun in glioblastoma by internal ribosome entry site (IRES)-mediated translational activation. *Proceedings of the National Academy of Sciences of the United States of America* **2012,** *109* (42), E2875-84.
277. Eferl, R.; Ricci, R.; Kenner, L.; Zenz, R.; David, J. P.; Rath, M.; Wagner, E. F., Liver tumor development. c-Jun antagonizes the proapoptotic activity of p53. *Cell* **2003,** *112* (2), 181-92.
278. Langer, S.; Singer, C. F.; Hudelist, G.; Dampier, B.; Kaserer, K.; Vinatzer, U.; Pehamberger, H.; Zielinski, C.; Kubista, E.; Schreibner, M., Jun and Fos family protein expression in human breast cancer: correlation of protein expression and clinicopathological parameters. *European journal of gynaecological oncology* **2006,** *27* (4), 345-52.
279. Mariani, O.; Brennetot, C.; Coindre, J. M.; Gruel, N.; Ganem, C.; Delattre, O.; Stern, M. H.; Aurias, A., JUN oncogene amplification and overexpression block adipocytic differentiation in highly aggressive sarcomas. *Cancer cell* **2007,** *11* (4), 361-74.
280. Rangatia, J.; Vangala, R. K.; Singh, S. M.; Peer Zada, A. A.; Elsasser, A.; Kohlmann, A.; Haferlach, T.; Tenen, D. G.; Hiddemann, W.; Behre, G., Elevated c-Jun expression in acute myeloid leukemias inhibits C/EBPalpha DNA binding via leucine zipper domain interaction. *Oncogene* **2003,** *22* (30), 4760-4.
281. Matlashewski, G.; Lamb, P.; Pim, D.; Peacock, J.; Crawford, L.; Benchimol, S., Isolation and characterization of a human p53 cDNA clone: expression of the human p53 gene. *The EMBO journal* **1984,** *3* (13), 3257-62.
282. Surget, S.; Khoury, M. P.; Bourdon, J. C., Uncovering the role of p53 splice variants in human malignancy: a clinical perspective. *OncoTargets and therapy* **2013,** *7*, 57-68.
283. Oren, M., Decision making by p53: life, death and cancer. *Cell death and differentiation* **2003,** *10* (4), 431-42.
284. Taylor, W. R.; Stark, G. R., Regulation of the G2/M transition by p53. *Oncogene* **2001,** *20* (15), 1803-15.
285. Raetz, C. R.; Whitfield, C., Lipopolysaccharide endotoxins. *Annual review of biochemistry* **2002,** *71*, 635-700.
286. Wu, F.; Bui, K. C.; Buckley, S.; Warburton, D., Cell cycle-dependent expression of cyclin D1 and a 45 kD protein in human A549 lung carcinoma cells. *American journal of respiratory cell and molecular biology* **1994,** *10* (4), 437-47.
287. Hoetelmans, R.; van Slooten, H. J.; Keijzer, R.; Erkeland, S.; van de Velde, C. J.; Dierendonck, J. H., Bcl-2 and Bax proteins are present in interphase nuclei of mammalian cells. *Cell death and differentiation* **2000,** *7* (4), 384-92.
288. Zinzalla, G.; Haque, M. R.; Basu, B. P.; Anderson, J.; Kaye, S. L.; Haider, S.; Hasan, F.; Antonow, D.; Essex, S.; Rahman, K. M.; Palmer, J.; Morgenstern, D.; Wilderspin, A. F.; Neidle, S.; Thurston, D. E., A novel small-molecule inhibitor of IL-6 signalling. *Bioorg Med Chem Lett* **2010,** *20* (23), 7029-32.
289. Milagre, C. S.; Gopinathan, G.; Everitt, G.; Thompson, R. G.; Kulbe, H.; Zhong, H.; Hollingsworth, R. E.; Grose, R.; Bowtell, D. D.; Hochhauser, D.; Balkwill, F. R., Adaptive Upregulation of EGFR Limits Attenuation of Tumor Growth by Neutralizing IL6 Antibodies, with Implications for Combined Therapy in Ovarian Cancer. *Cancer research* **2015,** *75* (7), 1255-64.
290. Wajant, H.; Pfizenmaier, K.; Scheurich, P., Tumor necrosis factor signaling. *Cell death and differentiation* **2003,** *10* (1), 45-65.
291. Chang, L.; Karin, M., Mammalian MAP kinase signalling cascades. *Nature* **2001,** *410* (6824), 37-40.
292. Kong, L.; Wang, X.; Zhang, K.; Yuan, W.; Yang, Q.; Fan, J.; Wang, P.; Liu, Q., Gypenosides Synergistically Enhances the Anti-Tumor Effect of 5-Fluorouracil on Colorectal Cancer In Vitro and In Vivo: A Role for Oxidative Stress-Mediated DNA Damage and p53 Activation. *PloS one* **2015,** *10* (9), e0137888.

293. Ferrara, N.; Davis-Smyth, T., The biology of vascular endothelial growth factor. *Endocrine reviews* **1997,** *18* (1), 4-25.
294. Raynaud, F. I.; Orr, R. M.; Goddard, P. M.; Lacey, H. A.; Lancashire, H.; Judson, I. R.; Beck, T.; Bryan, B.; Cotter, F. E., Pharmacokinetics of G3139, a phosphorothioate oligodeoxynucleotide antisense to bcl-2, after intravenous administration or continuous subcutaneous infusion to mice. *The Journal of pharmacology and experimental therapeutics* **1997,** *281* (1), 420-7.
295. Grimley, P. M.; Fang, H.; Rui, H.; Petricoin, E. F., 3rd; Ray, S.; Dong, F.; Fields, K. H.; Hu, R.; Zoon, K. C.; Audet, S.; Beeler, J., Prolonged STAT1 activation related to the growth arrest of malignant lymphoma cells by interferon-alpha. *Blood* **1998,** *91* (8), 3017-27.
296. Huang, S.; Bucana, C. D.; Van Arsdall, M.; Fidler, I. J., Stat1 negatively regulates angiogenesis, tumorigenicity and metastasis of tumor cells. *Oncogene* **2002,** *21* (16), 2504-12.
297. Thurston, D. E., Chemistry and Pharmacology of Anticancer Drugs. *CRC Press (Taylor & Francis): Boca Raton, FL,* **2006,** *1*, 281.
298. Thurston, D. E., *Chemistry and pharmacology of anticancer drugs*. CRC Press/Taylor & Francis: Boca Raton, 2007; p 290 p.
299. Case, D. A., Darden, T. A., Cheatham III, T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Walker, R. C., Zhang, W., Merz, K. M., Roberts, B., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolossváry, I., Wong, K. F., Paesani, F., Vanicek, J., Liu, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, M.-J., Cui, G., Roe, D. R., Mathews, D. H., Seetin, M. G., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., Kollman, P. A., AMBER 11, University of California, San Francisco, 2010. 2010.
300. Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E., 3rd; Laughton, C. A.; Orozco, M., Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers. *Biophysical journal* **2007,** *92* (11), 3817-29.
301. Rao, S. N.; Singh, U. C.; Kollman, P. A., Molecular mechanics simulations on covalent complexes between anthramycin and B DNA. *J Med Chem* **1986,** *29* (12), 2484-92.
302. Humphrey, W.; Dalke, A.; Schulten, K., VMD: visual molecular dynamics. *J Mol Graph* **1996,** *14* (1), 33-8, 27-8.
303. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **2004,** *25* (13), 1605-12.